MEASURING STREET-LEVEL WALKABILITY THROUGH BIG IMAGE DATA AND ITS ASSOCIATIONS WITH WALKING BEHAVIOR

A Dissertation
Presented to
The Academic Faculty

By

Bon Woo Koo

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of City and Regional Planning

Georgia Institute of Technology
August 2021

MEASURING STREET-LEVEL WALKABILITY THROUGH BIG IMAGE DATA AND ITS ASSOCIATIONS WITH WALKING BEHAVIOR

Approved by:

Dr. Nisha Botchwey, Advisor
School of City and Regional Planning
Georgia Institute of Technology

Dr. Subhrajit Guhathakurta
School of City and Regional Planning
Georgia Institute of Technology

Dr. Omar Isaac Asensio
School of Public Policy
Georgia Institute of Technology

Dr. Aaron Hipp
Department of Parks, Recreation and Tourism Management
North Carolina State University

Date Approved: June 28, 2021

ACKNOWLEDGEMENTS

I am grateful for the opportunity to have two wonderful advisors, Dr. Subhrajit Guhathakurta and Dr. Nisha Botchwey. They provided the scholarly training and experience as advisors as well as support and guidance as mentors. I thank my committee members, Dr. Omar Isaac Asensio and Dr. Aaron Hipp, for their invaluable insights and feedback. I would like to thank the students, faculties, and staffs at the School of City and Regional Planning for instruction, care, and kindness. I thank my colleagues at the Center for Spatial Planning Analytics and Visualization for scholarly discussions and leisurely chats whenever I needed them.

I am thankful for my family for their encouragement to pursue my goal. I cannot thank my wife Hae Seung enough for her decision to leave where she called home to accompany me in a foreign country. I could not have finished this work without her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The built environment characteristics associated with walkability range from neighborhood-level urban form factors to street-level urban design factors. However, many existing walkability measures are primarily based on neighborhood-level factors and lack consideration for street-level factors. Neighborhood-level factors alone can be limited in representing various needs of pedestrians. While pedestrians seek to fulfill their needs for accessibility, safety, comfort, and pleasurability, neighborhood-level factors tend to be limited to capturing the accessibility of the built environment (i.e., having places to go to and being physically connected to those places). The high-order needs (i.e., safety from crime, comfort from vehicular traffic, and aesthetic pleasurability) can be more closely proxied by street-level factors. Also, past studies suggested that certain street-level factors may weaken (or strengthen) the effect of neighborhood-level factors on walking behavior, which can be particularly important for disadvantaged populations who tend to be less responsive to neighborhood-level factors. However, measuring street-level factors often requires extensive manual labor and tends to be resource-intensive, resulting in the omission of street-level factors in widely used walkability measures such as Walk Score.

This dissertation uses street view images and computer vision to overcome these challenges in measuring street-level factors and expands the literature by examining their association with walking mode choice. This dissertation first applies a pre-trained computer vision model to street view images and measure mesoscale (i.e., a midlevel spatial scale between macro and microscale) factors of walkability. It finds that the

mesoscale factors have a significant contribution to walking mode choice models, and the contribution is greater than that from neighborhood-level factors. Next, the dissertation develops a method for automatically auditing walkability factors in microscale (i.e., the smallest spatial scale that pertains to the most fine-grain design details and their qualities) using the combination of computer vision, street view images, and geographic information systems. The validation results demonstrate moderate to high reliability between audit results by automated audit method and a trained human auditor. Finally, the dissertation uses automatically audited microscale factors to unpack the reasons for the weaker relationship between neighborhood-level factors and disadvantaged populations' walking behavior. The result shows that microscale factors play a sizable role in moderating the effect of neighborhood-level factors.

Collectively, this dissertation demonstrates the potential of using street view images and computer vision for research on the built environment-walking relationship and for collecting data on street-level factors over expansive geographic areas, a task that has traditionally been prohibitively expensive. The theoretical and methodological contributions of this dissertation help urban planners and designers understand the physical condition of their cities at street-level and make targeted interventions that are effective and equitable.

CHAPTER 1. INTRODUCTION

It is widely accepted that urban residents' travel behavior is influenced by the built environment. With growing concerns about the lack of physical activity and the excessive use of automobiles, the characteristics of the built environment that encourage walking and other active modes of transportation have gained prominence. Over the years, the literature on the association of various built environmental factors on active transportation and physical activity has greatly expanded. Based on the social ecological model, the researchers in transportation and public health have identified various built environmental factors as important influencers to behaviors related to active transportation and physical activity. Among various forms of active transport and physical activity, walking has been considered one of the most accessible and fundamental components as it requires no special equipment or training.

Various walkability indices have been developed to objectively measure the characteristics of the built environment that are conducive to walking behaviors, ranging from indices that focus on neighborhood-level urban form factors (e.g., population density, land use diversity, and street connectivity) to those focusing on street-level urban design factors (e.g., the scale and proportion of streets, the design and condition of buildings, and street furniture). To date, the majority of such indices, particularly ones that have a broad geographic coverage such as Walk Score (Walk Score, n.d.) and the National Walkability Index (U.S. Environmental Protection Agency, 2015), are constructed with mostly neighborhood-level factors. Street-level factors are included to a limited extent due to the constraints in data availability (Harvey & Aultman-Hall, 2016).

Although walkability indices that are based on neighborhood-level factors have been generally proven to be effective (Chiu et al., 2015; Duncan et al., 2011; Manaugh & El-Geneidy, 2011), the hierarchy of walking needs hypothesis by Alfonzo (2005) suggests that neighborhood- and street-level factors can contribute differently to different levels of walking needs. For example, neighborhood-level factors can be more closely associated with the accessibility need in Alfonzo's hierarchy, which is a basic level of walking need (e.g., having places to go to and being functionally connected to those places). Street-level factors often are more closely linked with higher-level needs, such as the need for safety, comfort, and pleasurability (e.g., the quality of the experience going to places) (Adkins et al., 2012; Alfonzo, 2005). Street-level factors can be particularly important as a place can have walking-conducive neighborhood-level factors but have poor street-level factors (Bereitschaft, 2017; Zhu & Lee, 2008), and recent studies report the importance of street-level factors in walking behavior (Adkins et al., 2012; Ewing & Clemente, 2013; Foltête & Piombini, 2007; Gallimore et al., 2011).

One of the reasons for failing to incorporate the street-level factors in walkability indices is the difficulty in obtaining objective measurements for large geographic areas (Harvey et al., 2015). Traditionally, street-level factors have often been measured using methods such as audit tools, expert evaluations, or surveys of participants' perceptions. Although these methods have provided invaluable ways to operationalize walkable streetscapes, audit tools and expert evaluations in particular are usually resource-intensive and time-consuming, making it difficult to be scaled up to large areas such as a city or a region.

Recent advances in computer vision techniques and increasingly available street view image datasets offer a unique opportunity to address this limitation by allowing researchers to automatically quantify street-level factors in more scalable ways than the traditional methods. Although a few pioneering studies have tested these techniques along similar lines of research and reported some promising outcomes (Dubey et al., 2016; Glaeser et al., 2018; Li, Santi, et al., 2018; Seiferling et al., 2017; Tang & Long, 2018; Wang, Helbich, et al., 2019; Yin & Wang, 2016; Zhang et al., 2019), there remain many limitations. A few important limitations include that few studies focused on measuring walking-inducive built environment and how the measurement relates with walking behavior and that those that did focus on walkability used measures that are too simplistic to represent a multifaceted concept of walkability (e.g., using sky view factor to represent walkability).

This dissertation examines the relationship between both neighborhood- and street-level walkability factors and walking mode choice by incorporating novel methods for measuring street-level factors using street view images and computer vision. After covering the findings and gaps in the literature in the next chapter, Chapter 3 focuses on measuring walkability factors in mesoscale, a midlevel spatial scale in street-level that is smaller than neighborhood-level but larger than the most fine-grained details of streetscapes, and how mesoscale factors are associated with walking behaviors. Chapter 4 aims to develop an automated method for auditing walkability factors in microscale, the most fine-grained spatial scale in street-level, and validating the performance of the developed method. Chapter 5 expands the third chapter by incorporating microscale factors with a focus on examining how microscale factors moderate the relationship between macroscale factors of walkability and walking behavior. The dissertation concludes by discussing the contributions of this dissertation on various ends and providing policy implications of the findings.

CHAPTER 2. SUMMARY OF THE CURRENT KNOWLEDGE AND GAP

2.1 Built Environment Factors Influencing Walkability

The three 'D' variables (3D) - density, diversity, and design, or the five 'D' variables (5D) which add destination accessibility and the distance to transit to the 3Ds, have been the foundational framework of numerous studies on travel behavior and walkability (Ewing & Cervero, 2010; Smith et al., 2008). Similarly, Handy et al. (2002) list density and intensity of development, land use mix, street connectivity, the scale of streets, and aesthetic qualities of a place as the dimensions of the built environment influencing physical activity. The dimensions in these frameworks encompass various spatial scales, and some studies have grouped these dimensions into two broad categories; neighborhood- and street-level factors (Cain et al., 2014; Harvey & Aultman-Hall, 2016; Mertens et al., 2015).

Neighborhood-level factors consist of macro-scale characteristics such as residential density, land use diversity, distance to destinations, and street connectivity (Ewing & Clemente, 2013; Sallis et al., 2011). Density and diversity are aggregate characteristics of the built environment that are often defined and measured at some aerial units (e.g., Census Tracts) – hence the name 'neighborhood-level.' Density and diversity of different types of land uses contribute to walkability by placing more activities in a given land area and by mixing different types of origins and destinations in proximity (Saelens et al., 2003). Street connectivity relates to the directness of travel on the street network (Saelens et al., 2003). For example, even when the straight-line distance is the same, actual travel distance may be shorter when the street network follows a grid pattern than when streets are sparsely connected like the ones commonly found in low-density suburbs (Saelens et al., 2003).

Most of the existing walkability indices are constructed based on neighborhood-level factors. Neighborhood-level factors are commonly measured using population, housing or employment data, street network, land or building use, and business location data, which are usually more widely available than data for street-level factors. Walk Score, one of the most widely used walkability indices, is one such example. Walk Score calculates its score based on the walking route distance from a given address to potential walking destinations (Walk Score, n.d.). The calculation also includes pedestrian friendliness metrics such as population density, intersection density, and the average block length.

Street-level factors, in general, are the streetscapes and design details that are smaller in scale than neighborhood-level factors, such as the configuration of street width and building height, the style and material of buildings, street trees and other planters, pedestrian-friendly facades, and street furniture and other fixtures. In addition to the scale, another significant distinction between neighborhood-level and street-level factors is that the street-level factors are commonly defined at eye level and are more *visually perceivable* than neighborhood-level factors. Street-level factors are closely linked with the scale and aesthetic qualities of streets (Handy et al., 2002). These factors shape urban design qualities (e.g., imageability, enclosure, human scale, and transparency) that influence walkability through eliciting individuals' reactions such as a sense of safety, sense of comfort, and level of interest (Ewing & Handy, 2009). These reactions “contribute to an overall perception of walkability and, ultimately, walking behavior” (Adkins et al., 2012, p. 500).

Some studies suggest that street-level factors can further be divided into two subgroups: (1) the structural form of streets and (2) finer design details attached to the structural form (Handy et al., 2002; Harvey & Aultman-Hall, 2016). The structural form determines “three-dimensional space along a street as bounded by buildings or other features (e.g., trees or walls)” (Handy et al., 2002, p. 66). Similarly, Harvey & Aultman-Hall (2016) define streetscapes as “the size and arrangement of large objects such as buildings and trees” (p. 149) and proposes to call it *mesoscale*, a midlevel spatial scale between macro- and micro-scale. The most fine-grained design details, such as memorable details on buildings and pedestrian-friendly façade, are considered microscale. This layer functions like a skin covering the structural form determined by the mesoscale streetscapes (Harvey & Aultman-Hall, 2016). Because microscale design details often involve highly granular features, automatically measuring them can be challenging (Harvey et al., 2017).

2.2 Hierarchy of Walking Needs Hypothesis

Considering both neighborhood- and street-level factors in measuring walkability can be important because neighborhood-level factors alone can be limited in representing various stages of individuals’ decision-making process for walking. Alfonzo’s hypothesis about the hierarchy of walking needs suggests that the most fundamental walking need is *feasibility*, which is the condition of individuals that makes a walking trip feasible, such as age, physical condition, or available time (Alfonzo, 2005). The next fundamental level of the hierarchy is *accessibility* (e.g., having places to go to and being functionally connected to those places). It is followed by other higher-order needs such as *safety, comfort,* and *pleasurability* (e.g., the quality of the experience going to places). Note that *safety* here pertains closely to safety from crime and incivility while *comfort* is linked with safety from traffic. Pedestrians seek to fulfill the need for accessibility, safety, comfort, and pleasurability when they make the decision to walk (Alfonzo, 2005). These needs would be fulfilled if the characteristics of the built environment offer desirable accessibility, safety, comfort, and pleasurability to pedestrians who are considering walking in it. Therefore, accessibility, safety, comfort, and pleasurability can be considered both as the needs of pedestrians or the quality of the built environment.

Although there is not a clear distinction in terms of which scales of walkability factors are associated with which level of walking needs, past studies seem to suggest that the accessibility needs can be captured by macroscale (i.e., neighborhood-level) factors while higher-order needs can be more closely proxied by meso- or microscale factors (Adkins et al., 2012; Alfonzo, 2005; Harvey et al., 2015). For example, the accessibility needs of Alfonzo's hypothesis can be operationalized as the distance, the number, and the mix of destinations, and the completeness of walking infrastructure (Alfonzo, 2005). The three higher-order needs, on the other hand, may be operationalized using measures such as the presence of varied streetscapes and public spaces, the width of streets and sidewalks, the existence of sidewalk buffers, street trees, and medians, and street furniture and water fountains (Alfonzo, 2005). Similarly, comfort and safety are associated with the cross-sectional proportion of mesoscale streetscape, the number of buildings per 100m, and tree coverage (Harvey et al., 2015; Harvey & Aultman-Hall, 2015).

Alfonzo's hypothesis poses that "an individual would not typically consider a higher-order need in his or her decision to walk if a more basic need was not already satisfied" (Alfonzo, 2005, p. 818). From this hierarchical order, it can be conjectured that macroscale factors may have a more fundamental relationship with walking behavior than meso- or microscale factors. Some of the past findings seem to support Alfonzo's hypothesis by showing that microscale factors (e.g., benches at bus stops) tend to have weaker impacts on travel behavior than macroscale factors such as land use mix (Cervero & Kockelman, 1997). However, recent studies report significant effects of some street-level factors (i.e., meso- or microscale factors) even after controlling for neighborhood-level factors (Adkins et al., 2012, p. 201; Cain et al., 2014; Ewing & Clemente, 2013). Importantly, Alfonzo et al. (2008) empirically examine their hypothesis by sequentially adding the measures of accessibility, safety, comfort, and pleasurability into regression models. They found that the measures of accessibility and safety were significantly associated with the number of walking trips for all purposes as well as for the number of destination walking trips (e.g., going to parks, stores, works, or schools). For recreation trips, only the safety measure was significantly associated with the trip frequency. These studies indicate that both neighborhood- and street-level factors may be contributing to walking behavior, with some studies reporting a higher significance of the street-level factors.

Furthermore, some studies have suggested that street-level factors may be one of the variables that can moderate the effect of neighborhood-level factors on walking behavior. Examples of the moderators of neighborhood-level factors include street-level factors that relate to the sense of safety and aesthetic quality of streets (Lovasi, Neckerman, et al., 2009) and socioeconomic or demographic status that can restrict the access to vehicles and impose limitations to the mode choice options (M. A. Alfonzo, 2005). For instance, even if someone is in a high-density area with mixed land uses and good connectivity, the person is still unlikely to walk if there are serious concerns about safety from crime and traffic or aesthetic disorderliness. For this reason, many studies on walkability paid attention to disadvantaged populations because of their tendency to have more undesirable street-level walkability factors (Bereitschaft, 2017; Neckerman et al., 2009; Sallis et al., 2011) and restricted access to vehicles.

2.3 Measuring Street-Level Factors

To overcome the difficulty of measuring street-level factors, around early 2010s there was an increasing number of studies paying attention to the street view images or other sources of images such as dashboard cameras as a way to 'virtually' audit the built environment (Badland et al., 2010; Brookfield & Tilley, 2016; Hipp et al., 2013; C. M. Kelly et al., 2013). These studies generally reported that virtual audits (i.e., an audit done by human through street view images from the Internet) reasonably replicated in-person audits (i.e., an audit done by human through physically visiting the street) at reduced time and cost. The literature comparing virtual audit and in-person audit reported that objectively identifiable and large objects were reliably audited in virtual audits (Badland et al., 2010; Clarke et al., 2010; Griew et al., 2013; Rundle et al., 2011), but temporally variable items (e.g., litters on the sidewalk) and small items (e.g., items with a size of a bag pack) were more challenging to virtually audit (Clarke et al., 2010; Rundle et al., 2011). Importantly, although virtual audits can eliminate the time needed to travel to target streets, the audit time required for each question in an audit tool was similar to in-person audit (Rundle et al., 2011). This long per-question audit time may preclude using virtual audits for measuring streetscapes for a large geographic area such as a city or a county, for example.

Recently, an increasing number of recent studies have measured streetscape characteristics using street view images and computer vision techniques for fully automated measurements and linked the measurements to various behavioral and health-related outcomes. Focusing on streetscapes and walkability and walking behavior, Yin and Wang (2016) used GSV images collected from 311 street blocks in Buffalo, New York, USA, and found a significant relationship between the proportion of sky (i.e., visual enclosure) and pedestrian count and Walk Score. Wang, Lu, et al. (2019) examined the relationship between neighborhood street walkability and the mental health of older residents in 45 residential neighborhoods in Haidian District in Beijing, China. Using the average of the proportion of sky from street view images taken from road segments within a 1000m-buffer around each study neighborhood as a proxy for neighborhood walkability (i.e., the lower the proportion of sky, the higher the walkability), they found a positive relationship between walkability and mental health. Similarly using sky visibility to represent enclosure and the visibility of street greenery as streetscape measurements, Li et al. (2018) studied the relationships among enclosure and street greenery and walking trip count in Boston, Massachusetts, USA, in various land use types. Wang, Liu, et al. (2019) used street view images to predict perceptions of neighborhood appearance (wealthy, safe, lively, beautiful, boring, and depressing), which were then regressed against the total time and intensity of physical activity. They found positive associations between physical activity with safe, lively, and beautiful appearances and negative associations with depressing and boring appearances of neighborhood environments. Nguyen et al. (2019) used GSV images to characterize the national built environment with the presence of highways, rural, and grassland and examined their association with various health outcomes at county and Census tract levels. They found associations between greater presence of highways and lower chronic diseases and premature mortality as well as between characteristics of rural areas and multiple adverse health outcomes, including obesity, physical inactivity, and premature mortality. Hankey et al. (2021) used street environment characteristics derived from GSV images as well as other public data sources (e.g., Census, Google Point of Interest) to predict pedestrian and bicycle counts. They found that the inclusion of street-level data improved prediction accuracy and that street-level data can be a useful alternative to Census data. Some studies focused on greenery at eye level and computed green view index (GVI) by calculating the proportion of green shown in street view images. It was found that GVI provides unique information that other conventional data sources do not capture (Larkin & Hystad, 2019; Li et al., 2015) and that street view-derived GVI is more closely associated with walking time than traditional greenery variables (Ki & Lee, 2021).

Although street view image-based measurements have opened a new possibility for measuring meso- and microscale built environmental factors and for enhancing our understanding of how they relate to health outcomes, there remain important research gaps. First, because there are only a limited number of studies in the walkability literature that used street view images and automatic measurement techniques such as computer vision, the effectiveness and generalizability of this emerging measurement technique are less understood. Second, the measurements incorporated in the literature are often too simplistic (e.g., using sky view factor to represent walkability) to accurately capture the complex, multi-level dimensions that collectively compose ‘walkability’. Third, studies that did incorporate various streetscape objects used pre-trained models that are limited to predicting only predetermined objects, and many objects that are relevant to pedestrian experience are excluded from consideration. For example, poorly maintained sidewalks that constitute trip hazard, presence of curb cuts and sidewalk buffers, or the maintenance quality of buildings, are seldom included in the list of detectable objects in pre-trained models although they are associated with active transportation or physical activity (Sallis et al., 2015). Fourth, because these studies often incorporated a limited set of neighborhood-level factors, or entirely excluded them in some cases, the empirical understanding of the relationship between neighborhood- and street-level factors and their relative contribution to walking behavior is limited. Fifth, the existing walkability studies that incorporated meso- or microscale walkability factors using street view images and computer vision techniques often failed to control for individual travelers’ characteristics (Li, Santi, et al., 2018; Yin & Wang, 2016). Given that the individual travelers’ characteristics can be closely associated with the most fundamental layer in the hierarchy of walking needs, that is, the feasibility for walking, it is critical that they are included in the walkability models as controls. Finally, few studies considered the fact that the effect of walkability, particularly of neighborhood-level factors, can be moderated by various factors. The omission of the moderating effect in the research framework can be particularly problematic for disadvantaged populations who tend to be associated with variables that can curve the benefit of walkability. These limitations collectively warrant the need for this dissertation.

2.4 Theoretical Framework

This dissertation posits a theoretical framework that synthesizes the empirical findings and discussions in the relevant literature. As shown in Figure 1, the neighborhood- and street-level walkability factors together determine overall walkability. The overall walkability is then combined with travelers' individual-level factors to influence the decision to walk. This theoretical framework can be viewed as a part of the ecological model of physical activity in Sallis et al. (2012). For consistency, this dissertation heretofore refers to the three scales of walkability measurements as macro-, meso-, and microscale factors. It is important to clarify that this dissertation does not directly measure the walking needs. The concept of "walking needs" is introduced in this dissertation as a theoretical link to illustrate why considering various scales of measurements can be important in explaining walking mode choice. This theoretical framework encompasses the following three chapters.

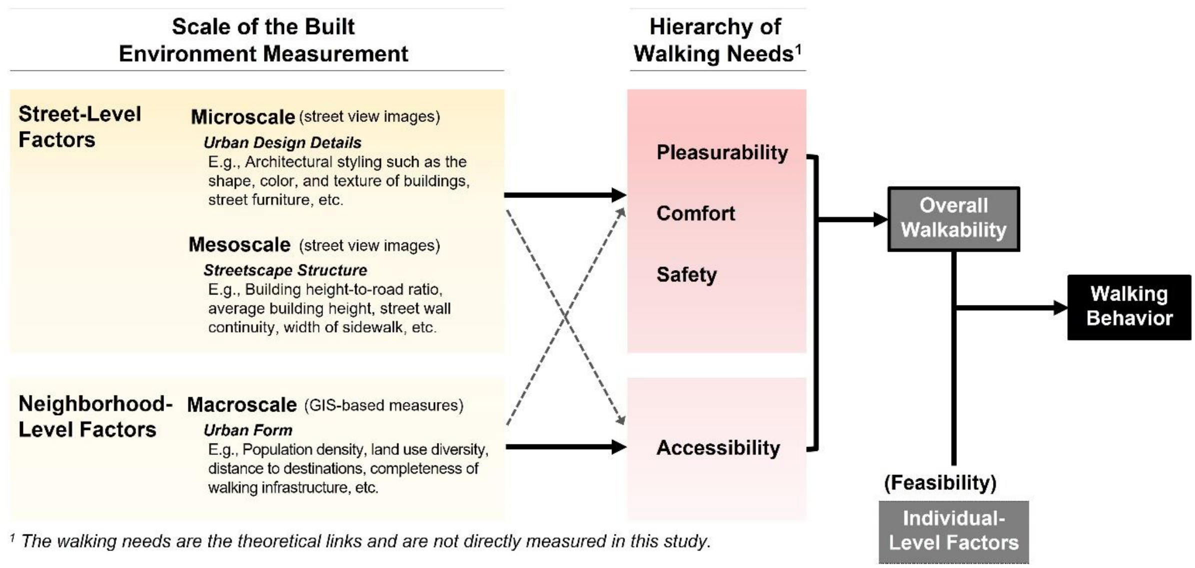


Figure 1. The Overall Theoretical Framework of the Analysis

CHAPTER 3. HOW ARE MACROSCALE AND MESOSCALE FACTORS ASSOCIATED WITH WALKING BEHAVIOR?

3.1 Introduction

The high cost of measuring street-level factors through the conventional methods led some scholars to focus on mesoscale factors as they are relatively small in scale and likely to be closely linked with higher-order walking needs (e.g., safety and comfort). Yet, they can be more amenable to objective and automated measurements than even smaller microscale factors (Harvey et al., 2017), as mesoscale streetscapes can be measured with metrics that define the boundary of the void between buildings and trees (Harvey et al., 2015, 2017; Harvey & Aultman-Hall, 2015).

Mesoscale factors are associated with the need for safety and comfort by providing the sense of *enclosure*, the urban design quality that "results when lines of sight are so decisively blocked as to make outdoor spaces seem room-like" (Ewing & Handy, 2009, p. 73). Harvey et al. (2015) used GIS and six mesoscale streetscape measurements to explain perceived safety in New York City, New York, USA. They found that street tree canopy, the number of buildings along a block, the cross-sectional proportion, and the length of a street segment have significant positive effects on perceived safety. They also found that the effects of mesoscale streetscape variables on perceived safety are greater than a neighborhood-scale urban form measure (in this study, Walk Score). Using the same set of mesoscale streetscape measurements, Harvey and Aultman-Hall (2015) investigated the relationship between the mesoscale streetscape and safety from traffic. After controlling for being on arterial roads, they found crashes that occurred on smaller, more enclosed streetscapes are less likely to be severe.

Although mesoscale factors are more amenable to objective measurements than microscale factors, there remain some limitations. First, GIS-based measures of mesoscale factors in the past studies are often measured at overhead-view and may not accurately represent street-level characteristics at eye-level. Because the measurements of the same item from overhead-view and eye-level view may not match well (Jiang et al., 2017), GIS-based measures can potentially deviate away from the perceptual stimuli that pedestrians receive at eye level while walking on streets. Second, mesoscale measures used in the literature require geospatial data from multiple sources, and some of them can be difficult to acquire from public data sources. For example, a detailed building height information required to compute cross-sectional proportion and average building height of a street, both of which are significant predictors of safety and comfort (Harvey et al., 2015; Harvey & Aultman-Hall, 2015), is often difficult to acquire except for a few major cities in the U.S.

This chapter measures mesoscale factors at eye level using more ubiquitously available street view images and a semantic segmentation technique and uses the measurements to examine the relationship between the built environment and walking mode choice. In doing so, this chapter addresses the limitations of the past studies by incorporating a comprehensive set of macroscale factors and multiple measures of mesoscale factors and controlling for personal as well as trip-level covariates.

3.2 Research Questions and Hypotheses

This chapter aims to answer the following questions: Is there a relationship between neighborhood- and street-level factors when evaluating the walkability of a place? To what extent does the inclusion of street-level or mesoscale factors improve our ability to explain people's walking mode choice? Based on the theoretical framework shown in Figure 1, these questions can be formally refined into hypotheses. The first hypothesis is that both macroscale and mesoscale factors will have statistically significant contributions to walking mode choice models when they are used separately. This hypothesis reflects the recent findings in the literature that challenge the past studies showing weak to insignificant associations of street-level factors on walking behavior. Second, it is hypothesized that macroscale and mesoscale factors will together explain the variation in walking behavior better than when they are used separately, given their unique and independent contributions. Note that the second hypothesis is, in essence, testing whether mesoscale factors, as measured using a computer vision technique, will incorporate particular information about the built environment that macroscale factors cannot. If, for example, mesoscale factors turn out to be highly correlated with macroscale factors, mesoscale factors would not have additional useful information about the walkability of the built environment compared to the macroscale factors.

3.3 Data and Analytical Methods

3.3.1 National Household Travel Survey data

All data used in this chapter were collected for the City of Atlanta, Georgia, USA. The travel data was extracted from the 2017 National Household Survey (NHTS). The

NHTS Georgia add-on data was provided by the Georgia Department of Transportation, from which the information about the location of the trip origin, the mode, purpose, and the travel distance of the trips, basic socioeconomic and demographic information, and other behavioral characteristics for each trip was acquired. Trips served as the unit of analysis. To retain a sufficient sample size while reducing the computation resources required to process GSV images, the analysis was limited to trips that have origins within the city boundary and did not consider the built environment of destinations. Following Cervero & Duncan (2003), the analysis was limited to trips "… that were unlikely to involve carrying significant amounts of items or goods, such as groceries" (9) and included trips that traveled for family/personal business, school or religious activities, socializing/recreational purposes, for eating out, or for shopping. The analysis was also narrowed to trips that traveled less than or equal to 1 mile, which is roughly the 90^{th} percentile of travel distance of walking trips in the data, as distances greater than that can quickly become challenging for walking. People who are less than 10-years of age and those who are using supportive devices such as wheelchair were removed from analysis, as their choice to walk can be affected by factors outside the scope of this chapter (e.g., the availability of caregivers). Item-nonresponses, missing values, and unrealistic data entries as determined by the data provider, the Federal Highway Administration, were excluded. Finally, travel mode was coded into a binary variable with labels walking and non-walking and used as the dependent variable.

The sociodemographic and other behavioral variables derived from NHTS include age, gender, race, educational attainment, number of vehicles owned by the household,

household income, driver status, number of walking activities in the past seven days, and the travel distance of each trip. These variables were included as control variables.

3.3.2 Google Street View Data

The image sampling method used in similar past studies can be categorized into two large groups: one group that focuses on capturing the built environment of intersection locations (e.g., Nguyen et al., 2019) and the other group that uses multiple images along the entire span of street segments with some fixed distance intervals such as 20, 50, or 100 meters (e.g., Ki & Lee, 2021). It is common for both methods to use 360º panoramic images for each image location, with some exceptions that relied on particular portions of images or directions (e.g., Yin et al., 2015). While focusing on intersections allows the measurement of streetscapes at important nodes of street networks with fewer images for a given areal unit, it is limited in capturing the streetscape characteristics that pedestrians experience as they move from one intersection to another. In contrast, while collecting images for the entire span of segments can provide a comprehensive view of the streetscapes, it requires a much greater number of images and computational resources for processing the images.

The GSV images used in this chapter were identified and downloaded using the following method, which combines the two approaches in the literature. First, four points were plotted for each street segment; two points at the middle of a street segment (heretofore, midpoints) and two points at the either ends of the street segments (heretofore, endpoints) using ArcGIS 10.5.1 and the road shapefile from Topologically Integrated Geographic Encoding and Referencing (TIGER) database. All points that are within 20-meters from expressway centerlines were deleted. Second, the heading of the camera was calculated differently for midpoints and endpoints. For midpoints, heading directions were calculated such that the sightline of the street view images is parallel to the street segment and that one image would be facing the opposite direction of the other image (i.e., looking back and forth towards the directions of walking; see Figure 2B and 2C). For endpoints, heading directions were calculated such that the sightline is parallel to the segment and that the image is looking into the street that is being measured. All other parameters were set to the default values. This procedure is applied to all street segments in the city to calculate the parameters needed to download street view images through GSV application programming interface (API). The total of 70,676 images with the size of 640 by 640 pixels were downloaded to cover the entire city area (see Figure 2A).

The NHTS was conducted between April 2016 and April 2017, and the metadata of GSV images showed that 92.1% of the images were taken between 2016 and 2018, indicating that nearly all of the street view images used in this chapter are temporally well-aligned with the NHTS data.

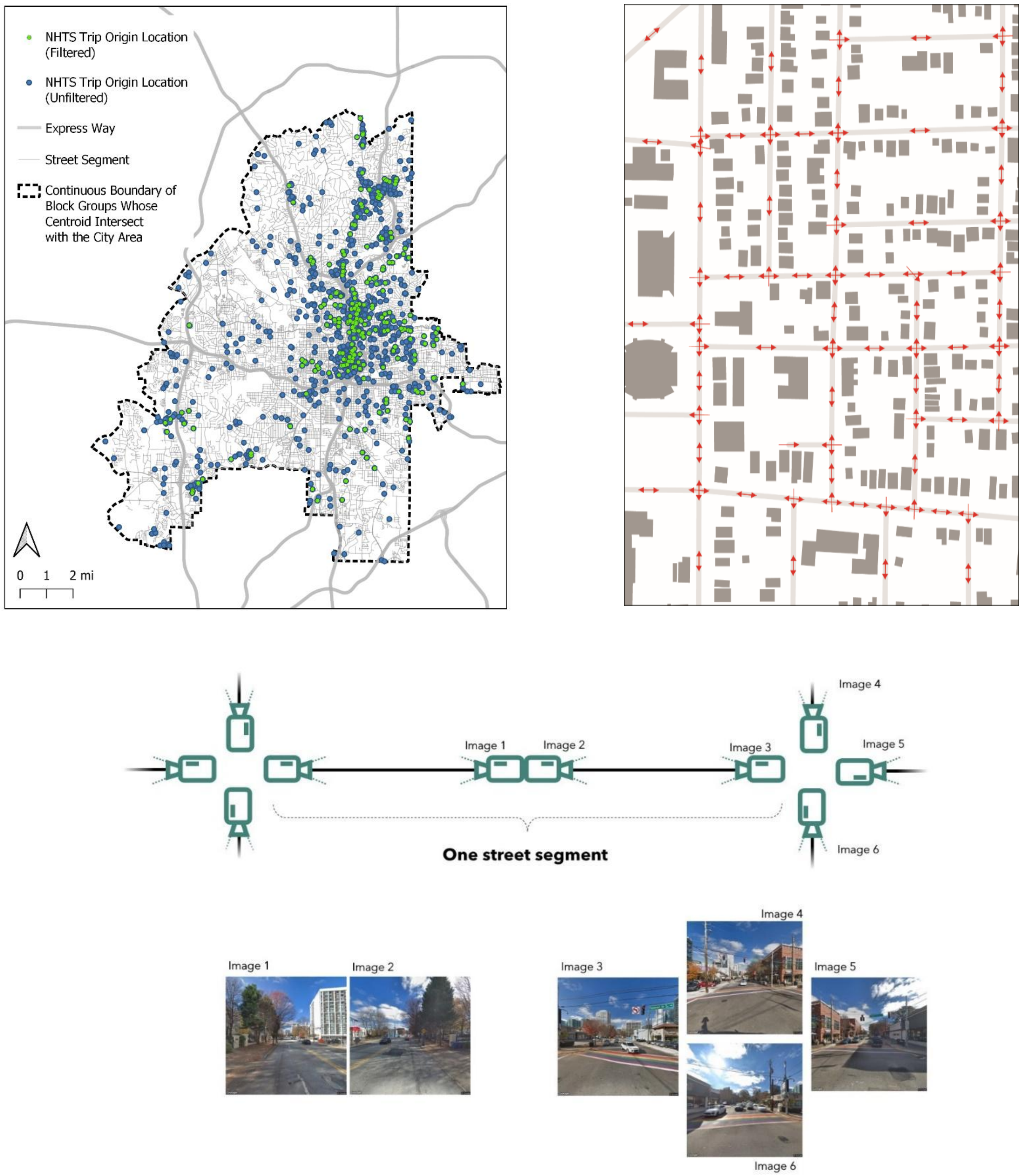


Figure 2. (A) The location of points where Google Street View images were downloaded, (B) the headings of the downloaded images, and (C) examples of downloaded Google Street View images

The collected images were processed through a semantic segmentation model called the Pyramid Scene Parsing Network (PSPNet) developed by Zhao et al. (2017). Built based on the fully convolutional network architecture (Long et al., 2015), a pre-trained PSPNet takes a raw image as an input, processes the image using pre-trained weights, and outputs a category with the highest probability as a prediction for each pixel. These weights are trained on ADE20K, a database that provides annotated images with 150 categories (Zhou et al., 2017). After the scene parsing through PSPNet was completed, the number of pixels per category in each image was counted. As shown in Figure 3, the seasonality did not appear to be a significant consideration as trees without leaves were correctly labeled.

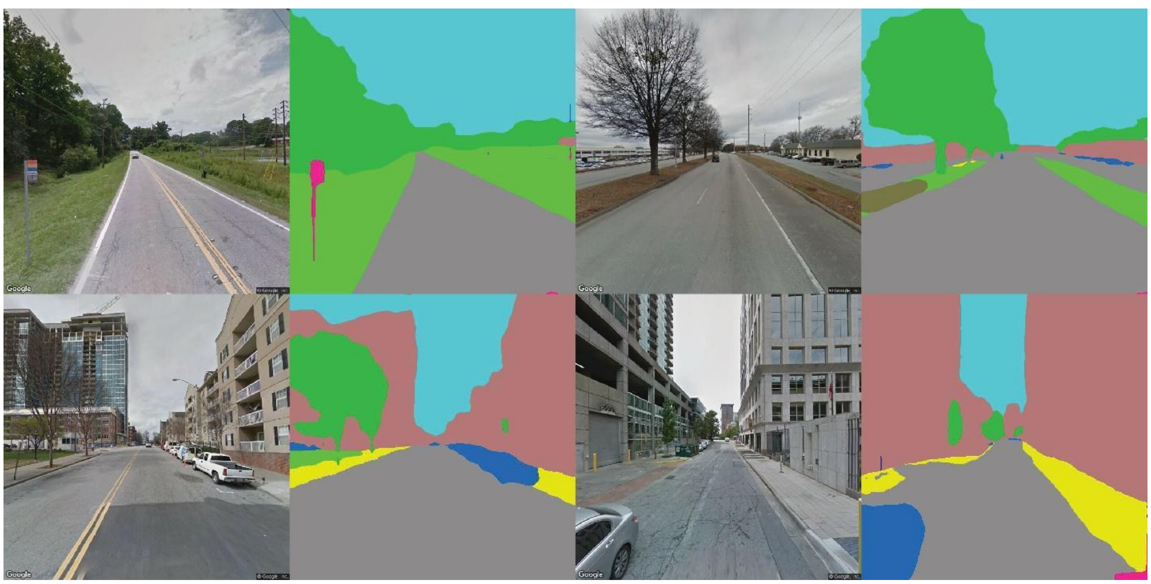


Figure 3. Examples of Google Street View images and their output from the computer vision processing by PSPNet

The output from PSPNet was examined to filter out any locations of anomalous images such as pictures of indoor spaces. This excluded 571 image locations from our dataset, leaving 70,105 image locations. Next, the number of pixels of each category shown in images taken at the same location but with different headings (directional) were averaged to represent the overall streetscapes of each location, reducing the number of image locations down to 31,351. These measures were then joined to the NHTS trip origins by drawing a buffer centered at each trip origin location and averaging the number of pixels of each category shown in images that fall in the buffer. The average length of street segments in the city of Atlanta is 148.8 meters, and 150-meter was used as the buffer distance. Operationally, the GSV images that are within about a block from the origin location were used to represent the streetscape of that specific origin location.

The final step was to select objects relevant to mesoscale streetscapes from the 150 categories and convert the averaged number of pixels of the selected categories described above into mesoscale measures of walkable streetscapes. Objects were selected if they are (1) consistent with the definition of mesoscale factors (e.g., bench, streetlight, and signboard are examples of excluded items for this reason), (2) usually found in outdoor spaces (e.g., wall, desk, and sofa were among those that were excluded for this reason), and (3) a part of the human-controlled environment (e.g., mountain and river, for example, are excluded for this reason; but landscape features such as trees and grass are included). The initial screening of potential mesoscale factor objects included building, tree, road, grass, sidewalk, plant, house, path, and skyscraper. Note that path and skyscraper were excluded as they were rarely detected. For example, over 75% of the NHTS trip origin points had nearly zero occurrences of path and skyscraper. Finally, vehicle categories are included in the consideration as they can block the view of the road on which they operate. Only car category was included as other vehicle categories such as bus and truck are rarely detected. A total of eight categories was selected to represent mesoscale streetscapes, including building, house, sidewalk, tree, road, grass, car, and plant. Detailed examples of each category are provided on the ADE20K website[1]. Based on the literature on urban design, built environment and active transport, and past studies on using street view images for measuring the built environment, the following three indices were formulated (Chen, 2017; Li et al., 2018; Tang & Long, 2018; Wang, Lu, et al., 2019; Yin & Wang, 2016; Zhang et al., 2019). The 'building-to street-ratio' was calculated as the ratio of the proportion of buildings to the sum of the proportion of sidewalk, road, and car. Note that car is included in the denominator because when a car is shown in an image, it is likely to be blocking the view of the road on which it stands. The 'greenness' was computed as the sum of the proportion of tree, grass, and plant. The 'sidewalk-to-street proportion' was measured as the proportion of the share of sidewalk to the sum of the share of sidewalk, road, and car. Heretofore, these three variables are called streetscape factors (see below for equations).

$$building\text{-}to\text{-}street\ ratio = \frac{\%\ building\ pixels\ +\ \%\ house\ pixels}{\%\ sidewalk\ pixels + \%\ road\ pixels\ + \%\ car\ pixels}$$

$$greenness = \%\ tree\ pixels + \%\ grass\ pixels + \%\ plant\ pixels$$

$$sidewalk\text{-}to\text{-}street\ proportion = \frac{\%\ sidewalk\ pixels}{\%\ sidewalk\ pixels + \%\ road\ pixels + \%\ car\ pixels}$$

[1] https://groups.csail.mit.edu/vision/datasets/ADE20K/

3.3.3 Macroscale Factors

Macroscale factors were selected and measured based on the 5D framework (Ewing & Cervero, 2010), which includes density, diversity, design, destination accessibility, and distance to transit. To represent the intensity of land uses, this chapter used employment density computed for a quarter-mile buffer around each NHTS trip origin instead of the commonly used residential density. One of the reasons the residential density was popularly used in the past studies examining the association between walkability and walking or physical activity is that the studies often focused on home-based trips (Cervero & Duncan, 2003; Frank et al., 2005; Manaugh & El-Geneidy, 2011; Park et al., 2015). Employment density was chosen for this chapter over residential density because only about 50% of the trips in the data after filtering are home-based trips and residential density may misrepresent the compactness of land uses in some cases. For example, the residential density of central business districts may not correctly reflect the intensity of daily pedestrian flow caused by the concentration of economic activities in the area. The 2015 employment data was downloaded from the Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES) at the Census Block scale. The employment density was computed as the number of jobs of all Census Blocks whose centroid intersects with the buffer divided by the area of the buffer. Diversity represents the degree to which different land uses are mixed in a given area, commonly measured using entropy indices. This chapter computed diversity using parcel-level land use data and the following formula:

$$diversity = 1 - \left(\frac{\sum_{i=1} \quad n_i(n_i - 1)}{N(N-1)}\right)$$

where n is the area of each land use category i in the a quarter-mile buffer around each NHTS trip origin (i = residential, commercial, institutional, and office uses); and N is the area of residential, commercial, institutional, and office uses combined.

The design in the 5D framework involves not only urban design details but also the design of street networks (Ewing & Cervero, 2010). Intersection density was computed as the number of all intersections falling into the quarter-mile buffer of NHTS trip origins divided by the area of the buffer. Destination accessibility was measured using Walk Score collected through the API for each location of NHTS trip origins. For Walk Score, any buffers were implemented because the construction of Walk Score already contains a similar mechanism in which the distance to nearby destinations is considered with a distance decay function (Walk Score, n.d.). Distance to transit was computed as the network distance from each NHTS origin to the nearest rail transit station in miles.

3.3.4 Analytical Methods

This chapter first used Pearson's correlation test to explore the relationship between macroscale and streetscape factors. Next, this chapter tested the two hypotheses by developing a series of logistic regression models. The decision to make a trip by walking is modeled using three binary logistic regression models. These models use trip mode (i.e., walking or non-walking) as the dependent variable. The Base Model includes only the control variables. Model 1 adds macroscale factors to the Base Model to examine how the macroscale factors are associated with the walking mode choice. Model 2 includes streetscape factors instead of macroscale factors. Model 3 includes all variables considered in this chapter. The macroscale factors include Walk Score, employment density, land use diversity, intersection density, and distance to transit. The streetscape factors include building-to-street ratio, greenness, and sidewalk-to-street proportion. Because the number of variables varies in different models, Adjusted McFadden's R^2 (adjusted R^2) and Bayesian Information Criteria (BIC) were selected to evaluate the model fit as these measures adjust for the number of variables included in the model.

3.4 Results

3.4.1 Descriptive Statistics

The NHTS data before the data filtration contained 2,189 trip records that have origin location was inside the study area. After the filtration as explained in Section 3.1, the data contained 329 trip records. Note that 11 NHTS origin locations did not have GSV images within 150-meters and were excluded from the analysis. The application of the 150-meter buffer reduced not only the NHTS origin locations but also the total number of GSV images used in the analysis down from 70,676 images to 8,149 images (roughly 11.5%) because there are many city areas where NHTS trip origin locations are too sparsely distributed and therefore many GSV images falling outside the 150-meter buffer.

In the final data used for the analysis, 204 trips (64.2%) were walking trips, and 114 trips (35.8%) were non-walking trips, totaling 318 trips. Because this filtering process removed trips that originated from outside the city area and those that traveled too far a distance to be walked, the final data leaned towards walking trips. Table 1 shows descriptive statistics of the variables used in the analysis.

Table 1. Descriptive statistics of the variables

Variable	Min	Median	Mean	Max	Std.dev.
Dependent variable					
Trip mode			Walking: 204 (64.2%) Non-walking: 114 (35.8%)		
Independent variable					
Age	11.0	40.0	42.7	85.0	17.0
Gender			Female: 134 (42.1%) Male: 184 (57.9%)		
Race			White: 209 (65.7%) African American: 88 (27.7%) Others: 21 (6.6%)		
Education			Less than high school: 11 (3.5%) High school or higher: 307 (96.5%)		
Count of household vehicles	0.0	2.0	1.6	6.0	1.0
Household income	5000.0	87499.5	96603.1	249998.0	71583.1
Driver status			Driver: 278 (87.4%) Non-driver: 40 (12.6%)		
Number walking trips in past 7 days	0.0	7.0	9.8	40.0	9.2
Travel distance (miles)	0.009	0.389	0.440	0.995	0.275
Distance from rail transit station	0.0	0.8	1.2	4.5	1.0
Employment density (count of jobs/km^2)	0.0	2822.0	11647.0	68623.6	17105.9
Intersection density (count of intersections/km^2)	10.0	53.7	60.1	159.2	29.2
Land use diversity	0.000	0.259	0.308	0.731	0.230
Destination accessibility (Walk Score)	7.0	81.5	76.1	98.0	18.4
Building-to-street ratio	0.0	0.2	0.3	1.1	0.3
Greenness	0.031	0.185	0.205	0.563	0.116
Sidewalk-to-street proportion	0.000	0.105	0.109	0.580	0.062

3.4.2 Correlation between Macro and Streetscape Factors

The correlation analysis showed strong correlations between the streetscape factors and macroscale factors in general, as shown in Table 2. An exception was the correlation coefficients among land use diversity and the streetscape factors, which recorded lower values that ranged between 0.068 to 0.437 in magnitude. Three macroscale factors – intersection density, distance to transit, and WalkScore –showed relatively consistent correlation coefficients with the three streetscape factors. In contrast, employment density and land use diversity showed varying levels of correlations with different streetscape factors. Employment density, for example, showed the strongest correlation across the board with building-to-street ratio with $r = 0.785$, while it was weakly correlated with sidewalk-to-street proportion with $r = 0.242$.

Additionally, the streetscape factors are correlated with each other. Building-to-street ratio is positively correlated with sidewalk-to-street proportion ($r = 0.439$, $p < 0.001$) but negatively associated with greenness ($r = -0.651$, $p < 0.001$). Greenness and sidewalk-to-street proportion have negative correlation ($r = -0.212$, $p < 0.001$). These negative correlations are not surprising as most of the tree canopy is located in low-density, single-family residential lots (Giarrusso & Smith, 2014) where sidewalks are often scarce. The variance inflation factor (VIF) was checked for all models in the subsequent analyses to ensure that multicollinearity does not severely bias the model results.

Table 2. Correlation between streetscape factors and macroscale factors

	Employment Density	Land Use Diversity	Intersection Density	Distance to Transit	Walk Score
Building-to-street ratio	0.785***	0.266***	0.681***	-0.611***	0.652***
Greenness	-0.512***	-0.437***	-0.474***	0.412***	-0.564***
Sidewalk-to-street proportion	0.242***	0.068	0.406***	-0.457***	0.433***

* Significant at the 5% level; **Significant at the 1% level; *** Significant at < 1% level.

3.4.3 The Relative Significance of the Macroscale and Streetscape Factors

The results from the binary logistic regressions are presented in Table 3. Standardized coefficients were generated for ease of comparison. Note that Table 3 presents the standardized coefficients in odds ratio form, reporting the odds of walking. No serious multicollinearity issue was found as the highest variation inflation factor across all models was 4.94. The control variables in the Base Model were generally significant at $\alpha = 0.05$ except for gender, race, and education. Among the person-level control variables, age, count of household vehicles, household income, driver status, and the number of walking activities in the past seven days were statistically significant, suggesting the importance of controlling for individual factors.

After accounting for the covariates, Model 1 showed that higher employment density and intersection density are positively and significantly associated with a greater odds of walking at $\alpha = 0.05$ and $\alpha = 0.1$, respectively, offering a substantial improvement in model fit over a model with only control variables (the likelihood ratio test between Base Model and Model 1: $\chi^2(5) = 33.057$, $p < 0.001$).

Model 2, which included the streetscape factors instead of the macroscale factors, provided a better model fit than Model 1 both in terms of adjusted R^2 and BIC. It showed that building-to-street ratio had the largest odds ratio and z-value among all built environment variables, making it one of the most significant predictors of a greater odds of walking. Similarly, greenness had a sizable odds ratio and was statistically significant, suggesting a positive association with a greater odds of walking. Sidewalk-to-street proportion was not significant.

Model 3 showed that when all available variables enter the model, the model fit was worse compared to Model 2 as measured by adjusted R^2 and BIC. Land use diversity was the only significant variable among the macroscale factors while building-to-street ratio and greenness remained significant. This result suggests that streetscape factors make a unique and independent contribution to explaining walking mode choice even after controlling for other macroscale factors. The significance of this contribution is also shown in a likelihood test result comparing Model 1 and Model 3 ($\chi^2(3) = 12.930$, $p = 0.005$). Furthermore, the comparison between Model 2 and Model 3 revealed that once the streetscape factors are included, macroscale factors do not seem to provide substantive improvements in the model fit. This finding is corroborated by a likelihood ratio test result comparing Model 2 and Model 3 ($\chi^2(5) = 8.346$, $p = 0.138$). Note that although the VIF values were not alarmingly high, the high correlations between some macro- and mesoscale factors could inflate the variance of the coefficient estimates in Model 3. To confirm that multicollinearity has not polluted the results, an investigation was done to examine how the results change when different pruning methods (e.g., various stepwise regression methods or excluding insignificant variables) are used. It was observed that the exclusion of variables reduced the VIF down to 2.087, but the relative importance of macroscale and streetscape factors in model fit did not change regardless of the method of variable selection. The relative size of the coefficient estimates and the statistical significance of macroscale and streetscape factors were also identical to Model 3 (these model results are available upon request).

Table 3. Results of the logistic regression models (dep.var = walking / non-walking in binary)

		Base Model	Model 1	Model 2	Model 3
	Constant	0.598	1.625	0.832	0.724
		(-0.566)	(0.524)	(-0.193)	(-0.316)
Personal-, trip-level covariates	Age	0.602**	0.688*	0.645*	0.640*
		(-2.973)	(-1.993)	(-2.293)	(-2.216)
	Gender – Male	1.229	1.013	1.183	1.168
	(Base: Female)	(0.648)	(0.037)	(0.474)	(0.424)
	Race – African American	0.781	0.400†	0.682	0.499
	(Base: white)	(-0.622)	(-1.855)	(-0.855)	(-1.361)
	Race – All other races	3.225†	1.042	1.870	1.020
	(Base: white)	(1.779)	(0.058)	(0.896)	(0.027)
	Education – High school or higher	3.260	2.043	3.355	4.401
	(Base: Less than high school)	(1.345)	(0.816)	(1.361)	(1.559)
	Count of household vehicles	0.642*	0.618*	0.624*	0.603*
		(-2.231)	(-2.134)	(-2.133)	(-2.240)
	Household income	1.714**	1.386	1.237	1.295
		(2.886)	(1.524)	(1.008)	(1.160)
	Driver status – Non-Driver	3.151†	5.204*	5.852*	7.521*
	(Base: Driver)	(1.683)	(2.193)	(2.371)	(2.427)
	Number walking activities	2.438***	2.540***	2.356***	2.542***
	in the past 7 days	(4.666)	(4.565)	(4.512)	(4.629)
	Trip distance	0.246***	0.286***	0.268***	0.251***
		(-7.670)	(-6.401)	(-6.423)	(-6.387)
Macroscale factors	Employment Density		2.645**		1.079
			(2.616)		(0.161)
	Land Use Diversity		1.292		1.758*
			(1.313)		(2.422)
	Intersection Density		1.516†		1.150
			(1.784)		(0.521)
	Distance to Transit		0.786		0.800
			(-0.907)		(-0.815)
	Walk Score		0.769		0.878
			(-1.043)		(-0.488)
Streetscape factors	Building-to-Street Ratio			5.533***	4.945**
				(4.954)	(3.058)
	Greenness			1.608*	2.070**
				(2.119)	(2.655)
	sidewalk-to-street proportion			0.867	0.791
				(-0.743)	(-1.181)
No. of observation		318	318	318	318
LL		-134.074	-117.545	-115.253	-111.080
Adj. McFadden's R^2		0.301	0.356	0.377	0.373
Bayesian Info. Criteria		331.530	327.283	311.176	331.640

- The regression results are in $\frac{Odds\ Ratio\ ***}{(z\text{-}statistic)}$ format, where the Odds Ratio is the exponent of the standardized coefficient from the logistic regression.
† Significant at the 10% level; * Significant at the 5% level; **Significant at the 1% level; *** Significant at < 1% level.

3.5 Discussion and Conclusion

Reflecting on the two hypotheses, the results offer many notable findings. The first hypothesis – that both macroscale and mesoscale factors will have comparable contributions to walking mode choice models when they are used separately– is supported by our data. Model 1 showed that macroscale walkability factors can add statistically significant improvements to the model fit compared to the model with only the control variables. Similarly, the streetscape factors contributed significantly to the model fit, providing a statistically significant improvement. One notable finding is that streetscape factors provided noticeably better model fits than that provided by macroscale factors. This result lends further support to the previous research that reported the importance of street-level walkability factors. The second hypothesis – that the macroscale and streetscape factors will have their unique contributions to the models when they are used together, improving the fit of the walking mode choice models more than when they are used separately – is not supported. The best model fit overall was observed not in a model that used both macro and mesoscale factors but in Model 2, which only contained the streetscape factors.

These findings warrant an important question: Why did the streetscape factors derived from street view images and a computer vision technique perform better than macroscale factors in explaining walking behavior? Although future research is needed to better answer this question, some possible explanations can be drawn from past studies and our correlation analysis. First, street view images may provide a more accurate representation of the actual built environment that pedestrians experience. Street view images taken from the pedestrians' perspective can correctly represent the complex interactions of large street objects that jointly create the visual stimuli of pedestrians, which 2-dimensional GIS-based methods are limited in representing. Taking street trees as an example, even when the street tree coverage measured from overhead-view is similar, its visual dominance at pedestrian perspective can vary depending on various conditions, such as the height of trees and buildings that are lined with the trees, the vertical shape of its crown, other large objects blocking the view, and the curve and slope of the street segment. Jiang et al. (2017) found that remotely sensed tree cover density does not match well with eye-level measures of tree cover density except when tree cover is very sparse. Using overhead-view and eye-level view measures in Hong Kong, Lu et al. (2019) found that cycling behavior is positively associated with eye-level street greenness but not with overhead-view greenness. Similarly, a study conducted in Seoul, Korea, found that green view index derived from street view images at eye level is more closely associated with walking time than the traditional greenery variables (Ki & Lee, 2021).

Second, the moderate to high correlations between streetscape factors and some of the macroscale factors indicate that the image-based measurement may be good proxies for these factors. The correlation analysis showed that employment density and intersection density, the two macroscale factors that had significant associations with the odds of walking in Model 1 (i.e., the model that contained all macroscale factor and no streetscape factors), are significantly correlated with streetscape factors, particularly building-to-street ratio and greenness. These correlations make sense because high employment density translates to large, and often tall, buildings to accommodate jobs, resulting in high building-to-street ratios. The high demand for development, particularly in the city center, can lead to a lack of potential spaces for trees and urban vegetation (Koo et al., 2019), elevating the negative correlation between employment density and greenness. Also, such locations in Atlanta often have grid-like street patterns with ample intersections. In contrast, low-density locations that are highly dependent on automobiles often have low buildings with wide and curvy roads for vehicular traffic, resulting in low building-to-street ratios. The ability for streetscape factors to capture some important macroscale factors related to walkability is attractive because it allows us to rely on fewer variables that can be derived from street view images only.

The five walking needs and urban design qualities are introduced in this chapter as theoretical links between the built environment measurements and walking behaviors. While the exact mechanism through which the streetscape factors are linked with walking mode choice has several aspects that are outside the scope of this chapter, the coefficient estimates and the statistical significance of streetscape factors appear to be in alignment with the literature. A large building-to-street ratio and more greenness may provide a greater sense of enclosure, a sensation that makes street spaces seem like an outdoor room (Ewing & Handy, 2009; Harvey et al., 2017), which is formed chiefly due to sightlines being blocked by buildings and trees. Enclosure is discussed frequently in the literature to link the built environment, particularly mesoscale factors, with a greater perceived safety (Harvey et al., 2015) and comfort (Harvey & Aultman-Hall, 2015), more pedestrians on the streets (Yin & Wang, 2016), and better mental health of elderly (Wang, Lu, et al., 2019).

It is important to note that the statistical insignificance of many macroscale factors does not mean they are not important. The inclusion of macroscale factors in

Model 1 provided a significantly better model fit compared to Base Model. As mentioned in Cervero et al. (2009), the 5Ds are “overlapping Venn diagrams” (209), and it is suspected that the insignificance of individual macroscale factors is due to this overlap. In fact, when one macroscale factor at a time was entered into the model instead of using all five of them simultaneously in Model 1, all five macroscale factors had statistical significance at $\alpha = 0.01$ (these results are omitted for brevity).

This chapter has several limitations that could be the basis for future research. First, this chapter only considered trip origins due to limitations in data availability and computational resources. Second, because the spatial resolution of Walk Score is unknown, whether the Walk Score from the API is the score for the exact X and Y-coordinates of the trip origin is unknown as well. If the Walk Score database is calculated with some distance intervals and returns a score of the nearest data point, it is possible that the geographical mismatch between trip origins and the nearest Walk Score point may have biased the performance of Walk Score. Third, this chapter was limited by the design of the survey data which provided personal and trip-level variables. The limited sample size did not allow us to parse out our dataset based on more detailed trip purposes and types. Some of the variables that are insignificant in the regression results may as well due to the limited sample size. The categorization in the survey dataset that merged similar trip purposes did not allow us to separate shopping trips for clothes from those for groceries, although they may involve carrying items that can be distinctly different in weight. As theoretically suggested by the hierarchy hypothesis and empirically presented by Manaugh and El-Geneidy (2011), different walkability indices that are built on different walkability factors can have varying effectiveness depending on trip purposes or types. These limitations can be addressed in future studies by using primary data collection methods, as opposed to analyzing secondary data such as NHTS, as self-collected data can be designed for specific research purposes and designs. For example, some studies using self-collected data measured streetscape characteristics of the actual trip routes (e.g., Park et al., 2015). Considering that the streetscapes can have a larger variation even within a relatively small area, measuring streetscape factors at route-level may better capture the experience of travelers than measuring them using buffers and may result in even more substantial improvements to walking mode choice models. As the NHTS does not provide the exact route of travel, the streetscapes could not be measured at the route-level. Fourth, because of the limitations in the capabilities of the computer vision technologies at the time of this chapter, the measures of mesoscale factors incorporated in this chapter have room for improvements. Future studies may build on these findings for more refined measures as more advanced technologies become available. Similarly, the availability, image age, and variance of image age of street view images tend to be associated with socioeconomic status (Fry et al., 2020). Although this chapter provided evidence that these issues are not severe, they may have introduced subtle biases to the result of this chapter. Fifth, this study was conducted in a single city of Atlanta, which is known for its low-density development pattern. This unique development pattern may have influenced the results, and the generalizability of the finding is unknown. Future studies will benefit from incorporating a more diverse urban environment across multiple cities and regions. Importantly, this chapter has potential sources of biases such as selective mobility (i.e., the tendency of people to sort themselves into different neighborhoods to live or places to visits throughout a day based on their socioeconomic or other relevant status) and uncertain geographic context problem (i.e., a bias coming from using some arbitrary areal units for analyses due to the lack of knowledge about the precise ways in which the environment influences people's behavior) (Kwan, 2018, p. 1486). The degree to which the results of this chapter may be biased due to these issues is unknown. The potential biases call for future research with a more robust study design.

Due to its pedestrian perspective, wide coverages, and fine-grained spatial resolution, street view images provide a unique opportunity to refine the ways in which walkable environments are measured and advance our understanding of how the streetscape is linked with walking. In this chapter, GSV images were used to quantify the mesoscale attributes of streetscapes. This chapter examined the relationship between mesoscale and macroscale walkability factors and whether a computer vision-based measure of the streetscapes would contribute to explaining walking mode choice in addition to macroscale factors. The streetscape factors showed better model fits than macroscale factors, suggesting that the streetscape factors appear to be substitutes of macroscale factors rather than work synergistically with them. This chapter provided potential explanations for this result: the image-based streetscape factors as presented in this chapter may perform as proxies of some macroscale factors to some degree while providing the benefit of better representing pedestrian experience from an eye-level view.

CHAPTER 4. DEVELOPMENT AND VALIDATION OF AUTOMATED MICROSCALE WALKABILITY AUDIT METHOD

4.1 Introduction

The previous chapter used street view images and a pre-trained computer vision model called PSPNet and focused on mesoscale factors of walkability while controlling for macroscale factors, personal-, and trip-related covariates. This chapter examines the last component, the microscale factors of walkability.

Recently, microscale factors have gained attention among researchers as their importance on walking behavior has gained increasing support in the literature. Many walkability audit tools that focus on microscale factors have been developed and validated. For example, the Microscale Audit of Pedestrian Streetscapes (MAPS) has included fine-grained design details and maintenance quality of streetscape components. Studies using MAPS have reported significant associations with various types of walking and active transport and satisfaction on walking accessibility (Cain et al., 2014; Sallis et al., 2015). Furthermore, microscale factors are relatively easy, quick, and inexpensive to modify, making timely interventions for promoting active transport and physical activity much more feasible than macroscale factors.

Despite these strengths, microscale factors have been rarely incorporated into widely used walkability indices such as Walk Score because its measurements have heavily relied on on-site, manual audits or surveys. With the introduction of street view image services such as Google Street View, many studies examined the possibility of replicating in-person audits with virtual audits (i.e., audits that are done by human auditors using the Google Street View (GSV) service to reduce resources required to travel to the target streets). These studies generally reported high agreement between in-person and virtual audits and suggested that virtual audit using GSV images is a reliable method for auditing the built and social environment (Clarke et al., 2010; Griew et al., 2013; Gullón et al., 2015; Rundle et al., 2011).

However, although virtual audit can eliminate the time and resources required for in-person auditors to travel to the target streets physically, they still require a similar amount of audit time per item (Rundle et al., 2011) or even more time than in-person audit (Gullón et al., 2015). Assuming 10-15 minutes of audit time for each street segment (Griew et al., 2013), virtually auditing all 21,000 streets in the city of Atlanta is still prohibitively expensive and labor-intensive. This limited scalability of virtual audit indicates that virtual audits may not save enough time and resources to make the inclusion of microscale factors into widely used walkability indices feasible. Even with virtual audits, the identification of streets that need planning and design interventions remains challenging, and the geographic and temporal scope of research is limited.

Recent studies have shown that computer vision can efficiently and reliably extract the physical characteristics of streetscapes from street view images (Ki & Lee, 2021; Li, Ratti, et al., 2018; Li, Santi, et al., 2018; Tang & Long, 2018; Wang, Lu, et al., 2019). Using computer vision techniques, these studies have extracted some key qualities of streetscapes, such as how much greenery, building, and sky can be seen from street view images. These automatically extracted data have been used to illustrate the association between microscale built environment characteristics and outcomes of interest (e.g., mental health and human walking activity). However, the methods are insufficient in capturing the multidimensional attributes that mediate walkability because they are limited to extracting only one or a few of the factors that jointly contribute to walking behavior and are incapable of detecting the quality aspect of the environment.

This chapter presents one of the first studies to explore the feasibility of developing a method for automatically replicating the complete set of items of a validated walkability audit tool using computer vision and street view images. An important part of this study is the validation of the tool developed, which is undertaken by comparing the results from the automated audit with a virtual audit done by a trained human auditor.

4.2. Literature Review

Around 2010, about three years after the launch of the Google Street View image service, researchers in urban planning and public health started to examine the potential of GSV images in replacing in-person audits. The majority of these studies focused either on evaluating the agreement between the in-person audit and virtual audit or on the inter-rater reliability between virtual audits. The findings from the studies comparing in-person and virtual audits generally reported that the street view images could reliably replicate the in-person audit with some caveats. Clarke et al. (2010) compared in-person audit and virtual audit of 244 streets in Chicago, IL using a 29-item audit tool. They reported that indicators of "the built environment and neighborhood social and physical disorder" were reliably audited using the street view images. The observed agreement and kappa statistic (κ) were particularly high for objectively observable items such as "signs advertising alcohol (observed agreement = .92, κ = .34) or the presence of trees lining the street (observed agreement =.94, κ =.49)" (p.1227). However, the agreement was lower for items that need finer observation, such as the presence of garbage, litter, or broken glass (observed agreement=.35, κ =.04). They explained that there is a five-year time difference between when the in-person audit was conducted and the time the virtual audit was done, and these items are "... likely to have changed substantially over the five years between the in-person and virtual audit" (p. 1227). They noted the agreement between the in-person and virtual audit was comparable to the inter-rater reliability between in-person audits. They conclude that "... some of the variability in characteristics observed across modes of observation may in fact be due to inter-rater reliability or test-retest reliability over the five years between observations" (p. 1228).

Other studies generally shared the finding that virtual audits can reliably replicate in-person audits except for ephemeral items. Badland et al. (2010) conducted a similar study in New Zealand, comparing in-person and virtual audits on 48 streets using a 21-item audit tool. They found that in-person and virtual audit showed acceptable levels of agreement on the majority of the items with intraclass correlation coefficient $\geq$ .70. A few items, including the number of fixed traffic controls, neighborhood permeability, and land use mix, showed agreements below the acceptable agreement level. Similar to Clarke et al. (2010), there were about two years of time lag between the in-person and virtual audit, which may have contributed to the low agreement for ephemeral items (e.g., litters on sidewalks). Rundle et al. (2011) conducted a similar comparison between in-person and virtual audit on 38 high-walkability block segments in New York City, NY. Among 103 categorical measures, 82.5% of the items show moderate or higher agreement (i.e., agreement ≥ .60). Among 37 count or proportion items, 62.1% of the items showed moderate or higher correlation (i.e., Spearman rank-order correlation ≥ 0.40). The agreement and correlations were higher for large or less temporally variable items. One important note in Rundle et al. (2011) is that because most street view images are taken by cameras attached on top of cars, there usually is some distance between the location of cameras and the sidewalk. Small items can be difficult to discern, particularly when they are located on the sidewalk surface, low to the ground, as they can easily be hidden behind other objects. Another study done by Griew et al. (2013) showed a similar result in the UK, finding that "percent agreement between in-person and desk-based audits … was high across all street characteristic categories with results ranging from 75 to 97% agreement (average 84%) and the kappa co-efficient ranging from k = 0.5 to 0.9 (moderate to almost perfect)" (p.5). They noted that the inter-rater reliability varied substantially between land uses, with the lowest agreement found in industrial areas and the highest found in residential areas. Similar to all other studies, they explained that items requiring "a judgment on quality or aesthetics" are commonly found to have low reliability.

The next group of studies focused on inter-rater reliability of virtual audits (i.e., no in-person audits involved in the studies). Using an 89-item audit tool on 288 streets in Indianapolis and St. Louis, they reported high levels of mean the prevalence-adjusted bias-adjusted kappa statistic (PABAK) across all items (PABAK = 0.84) with 95% of the items having a substantial or near-perfect agreement.

The audit measurements using street view images are associated with various other outcome variables linked with the built environment condition. The outcome variables shown to have associations include observed physical activity (Kelly et al., 2014), children's antisocial behavior (Odgers et al., 2012), the severity of pedestrian crashes (Hanson et al., 2013), gentrification (Hwang & Sampson, 2014), neighborhood disorder (Bader et al., 2017; Marco et al., 2017), and violent crime (L. He et al., 2017). These studies further support the validity of using street view images for auditing the built as well as social environment.

The virtual audits used in most of these studies were, however, still manually done. For a fully automated extraction of built environmental information from street view images, many recent studies started utilizing emerging computer vision algorithms for detecting various streetscape objects. Summaries of these studies are discussed in detail in Chapter 2. Although this approach has proven to be highly effective in capturing some of the pedestrian-related streetscape objects, it is often limited by what the pre-trained, off-the-shelf computer vision models offer. While walkability is a composite concept that involves various factors from macroscale to microscale, many studies are limited to a few proxies of walkability (e.g., green view index or sky view factor) and rarely included microscale factors.

To the best of the author's knowledge, few studies, if any, have used fully automated methods (e.g., GIS or computer vision) to audit the full suite of items in validated walkability audit tools such as the Microscale Audit of Pedestrian Streetscapes (MAPS) or the Neighborhood Environment Walkability Survey (NEWS). The lack of an automated audit method suggests that, although street view images are reliable sources of data on microscale walkability factors and have a wide geographic and temporal coverage, prior studies have not taken full advantage of their benefits. The absence of an automated audit method can result in research and planning practices that are limited to either small geographic/temporal scope or incomplete sets of items that may lead to weak construct validity.

4.3 Data and Analytical Methods

Reliably replicating virtual audit with automated audit by computer vision and GIS is a multi-stage process. These stages include the selection of audit tools, the selection and training of computer vision models, the collection and processing of street view images, the training of a human auditor, and assessing the agreement of audits by the human auditor and computer vision models.

4.3.1 Audit tool

This chapter uses a shortened version of the Microscale Audit of Pedestrian Streetscapes (MAPS-mini) for the comparison between manual and automated audits. The MAPS-mini is a 15-item version of the full, 120-item MAPS tool developed by Sallis et al. (2015). This audit tool is chosen for this chapter because (1) it has been validated to have statistically significant associations with active transport such as walking and biking (Sallis et al., 2015), (2) its short design makes the development of training dataset for computer vision models more feasible under time and resource constraints, (3) the 15-item version maintains strong validity despite its short length, with a high correlation (r = .85) with its full length, 120-item MAPS, (4) it contains relatively fewer items that require subjective judgment compared to, for example, NEWS, and (5) the items included in the audit are relatively stable over time and do not tend to change rapidly.

The MAPS-mini consists of two parts: The Crossing and the Segment. As shown in Table 4 below, the Crossing part has questions on three items, including walk signals, curb ramps, and marked crosswalks. The Segment part contains the rest of the questions on street designs and qualities. The total point for each segment is calculated by summing the point for each item.

Some adjustments are made for a more efficient use of computer vision models. First, some items in the Segment part are audited using geographic information systems (GIS) rather than computer vision models, as they can easily, and perhaps more effectively, be audited using conventional GIS data. Many municipalities and government entities maintain GIS databases on land use and transportation infrastructures (e.g., bike facilities and transit stops), and this chapter uses such databases to audit the following items in MAPS-mini: the primary type of land use (i.e., residential or commercial?) and presence of parks, transit stops, and bike lanes. The second adjustment is that, while the original instruction for the MAPS-mini direct users to audit *only one of two crossings* of any given street segment, this chapter audits crossings on both ends of a segment separately. The original MAPS-mini is designed to be applied to all street segments in a given area or all street segments in a route between origin and destination. In such cases, auditing only one of two crossings can be sufficient, as it is guaranteed that the other crossing will be considered when the next street segment is audited. However, this chapter randomly sampled streets for validation, and the sampled streets are not contiguous, which requires that both crossings be audited. A modification corollary to this adjustment is that this chapter simplifies the curb ramp question (i.e., the second item in the Crossing part) by deleting the third option, "Yes, at both pre-crossing and post-crossing curb(s)." Because this chapter audits two crossings at both ends of a segment separately, this option is redundant. Third, while the streetlights in the MAPS-mini is categorical, this chapter records the count of streetlights because categorizing the count of streetlights into None, Some, and Ample was ambiguous, and creating an objective criterion for the categorization was challenging. As the presence of streetlights can be obtained post hoc from the count, this modification is not a loss of information.

Table 4.The original items on MAPS-mini and the method for automated audit. Total point is calculated by summing the scores in parenthesis across all items.

Part	**Item**	**Audit Method**
Crossing	Is a pedestrian walk signal present? *Possible answers: No(0) / Yes(1)*	Computer vision
	Is there a ramp at the curb(s)? Possible answers: No(0) / Yes (1)	
	Is there a marked crosswalk? *Possible answers: No(0) / Yes(1)*	
Segment	Type of land use? *Possible answers: Residential(0) / Commercial(1)*	GIS
	How many public parks are present? *Possible* answers*: 0(0) / 1(1) / 2 or more(2)*	
	How many public transit stops are present? *Possible answers: 0(0) / 1(1) / 2 or more(2)*	
	Is there a designated bike path? *Possible answers: No(0) / Painted line(1) / Physical Barrier(2)*	
	Are there any benches or places to sit (include bus stop benches)? *Possible* answers*: No(0) / Yes(1)*	Computer vision
	Are streetlights installed? *Possible* answers*: None(0) / Some(1) / Ample(2)*	
	Are the buildings well maintained? *Possible* answers*: 0-99%(0)/ 100%(1)*	
	Is graffiti/tagging present (do not include murals)? *Possible* answers*: No(0) / Yes(1)*	
	Is a sidewalk present? Possible *answers: No(0) / Yes(1)*	
	Are there poorly maintained sections of the sidewalk that constitute major trip hazards? (e.g., heaves, misalignment, cracks, overgrowth, incomplete sidewalk) *Possible answers: None(0) / Any or No sidewalk present(1)*	
	Is a buffer present? *Possible* answers*: No or No sidewalk present(0) / Yes(1)*	
	What percentage of the length of the sidewalk/walkway is covered by trees, awnings, or other overhead coverage? *Possible* answers*: 0-25% or no sidewalk(0) / 26-75%(1) / 76-100%(2)*	

4.3.2 Street View Images

For the accurate representation of the streetscapes, the street view images need to be systematically collected to ensure that (1) they cover all views of a given street segment that are relevant to the MAPS-mini and (2) that no two images have significant overlaps which would result in counting one object more than once. All street view images are downloaded through Google Street View API and are 640 by 640 pixels large. The field of view (FOV), a parameter that determines the horizontal field of view of the image, is set to 90 degrees.

This chapter collects images for the Crossing part and the Segment part of MAPS-mini separately. For each street segment (i.e., continuous stretch of a street defined by two intersection points at either end), there are two intersections, and two images are downloaded for each intersection, resulting in four intersection images for each street segment (heretofore, Crossing images). As there are many streets that are wider than what a street view image with 90 FOV can capture, two Crossing images are downloaded for one intersection, resulting in four Crossing images for one street segment. The headings of the Crossing images are calculated using the sf package in R 4.0.2. in the following ways: First, using the road centerline shapefile from the Topologically Integrated Geographic Encoding and Referencing database (TIGER), the heading from the first (last) vertex of a street segment to the second (second to the last) vertex. Call this heading a straight heading. Second, for each intersection, two images with headings equal to ‘straight heading – 45’ and ‘straight heading + 45’, respectively, are downloaded.

For the Segment part, the exact locations of Google Street View images are unknown a priori. First, the metadata of street view images is downloaded at every 5 meters to identify the exact locations of as many existing street view images as possible (see Figure 4). Once the locations of possible images are identified, the headings for each of those locations are calculated such that there are two images for each location that are perpendicular to the street segment and are looking back to back (i.e., looking at both sides of the segment). Finally, images looking up are downloaded for all of the image locations, creating a 'virtual tunnel' of images.

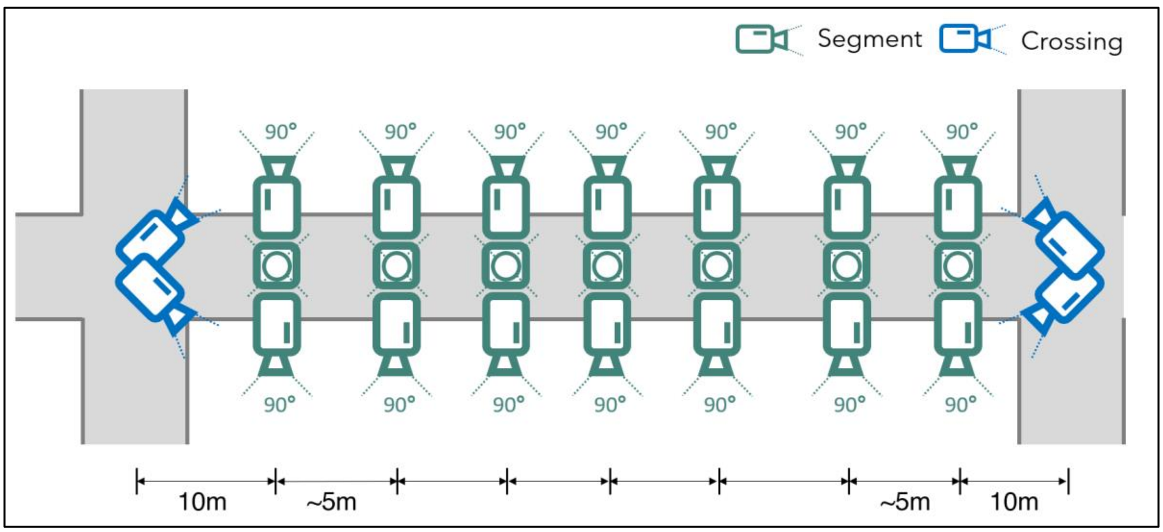


Figure 4. Location and heading of street view images.

The final processing of images involved cropping the overlaps or appending the gaps on the consecutive images. Note that objects that matter to pedestrian experience (e.g., sidewalks, buffers, and streetlights) are mostly found between buildings and the shoulders of the road. Holding the field of view constant, the longer the distance between the camera and the shoulder of a road (heretofore, this distance is called D), the longer the span of a street segment captured in an image. Assuming that the road is straight, if the distance between the location of one image and the next image (heretofore, this distance is called I) is shorter than 2D, objects that are at the shoulder of roads are likely to be captured in both images. This overlap is a non-issue for presence-based items (e.g., items such as 'is graffiti/tagging present?', which has two possible answers, 'Yes' or 'No') but can bias the results for count-based items (e.g., 'are streetlights installed?' with None, Some, or Ample answer). The opposite case of overlaps is gaps. This happens when I is longer than 2D. This can significantly affect both the presence-based and count-based items. To eliminate overlaps or gaps between consecutive image locations, this chapter develops a method that does the following tasks: (1) estimate the distance between the camera and the shoulder of the road, (2) calculate the proportion of the image that is estimated to be overlapped (or gapped) with the next image, and (3) crop out the overlapped portion or append the gap by downloading additional images with rotated camera heading. Detailed descriptions of each step are shown below.

4.4.3.1 Estimation of the distance between the camera and the road

This estimation uses the number of pixels that represent the road in an image. The information on the location of road pixels is acquired by applying Mask R-CNN, a computer vision model that will be described in the following subsection. Roads are usually shown at the bottom of street view images because the images are taken by a camera mounted on top of a car, which sits at about 8.2 feet[2] from the ground. The number of pixels representing the road is counted from the bottom of the image and

[2] https://petapixel.com/2012/10/15/a-glimpse-of-googles-fleet-of-camera-equipped-street-view-cars/#:~:text=Each%20car%20uses%2015%20cameras,a%20height%20of%208.2%20feet.

labeled as *n* (see Figure 5 below). The estimated distance (*d*) between the location of the camera and the shoulder of the road is calculated as

$$d = \frac{320}{320 - n} * 8.2$$

where 320 is the half of the height of street view images in pixels.

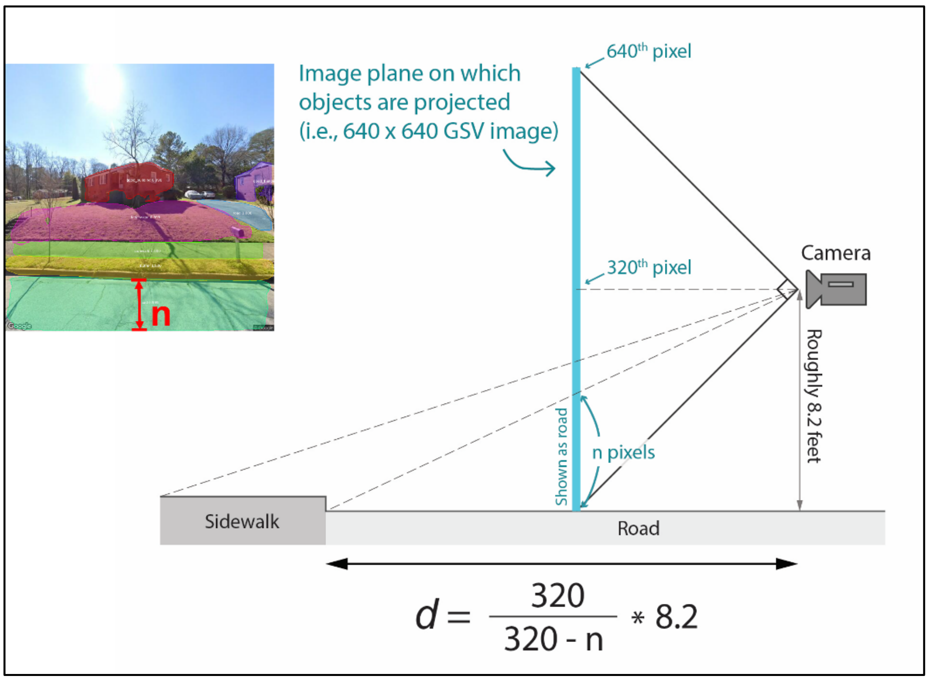


Figure 5. Method for estimating the distance from camera to the shoulder of road

4.4.3.2 Calculation of the proportion of the image that is estimated to be overlapped (or gapped) with the next image

This estimation uses the distance *d* and other information extracted from the metadata of street view images to calculate the proportion of overlapped portions between one image and the next image. This proportion can be defined as *x*/W where W is the width of the image plane at the location of the shoulder of the road measured in meters and calculated as 2*d*; *x* is the distance measured in meters on the image plane that is overlapped (or gapped) with the image plane of the next image (see Figure 6). The calculation of *x*/W requires other parameters, including I = the Euclidean distance between the location of one image (C_1) and the next image (C_2); and θ = the angle between C_1 and C_2. These parameters except *x* are known to the author, and the purpose here is to calculate *x*. The equations for calculating *x* differ depending on whether the images are looking into the curvature of the road or looking outward of the curvature of the road. When looking outward, *x* is calculated as

$$x = \frac{d * \sqrt{2} * cos\theta - \mathrm{I} * \sin(45° + \frac{\theta}{2})}{sin(135° - \theta)}$$

, and when looking inward, x is calculated using the following equation.

$$x = \frac{d * \sqrt{2} * cos^2\theta}{cos(45° + \theta)} - I * \sqrt{2} * sin(45° - \frac{\theta}{2})$$

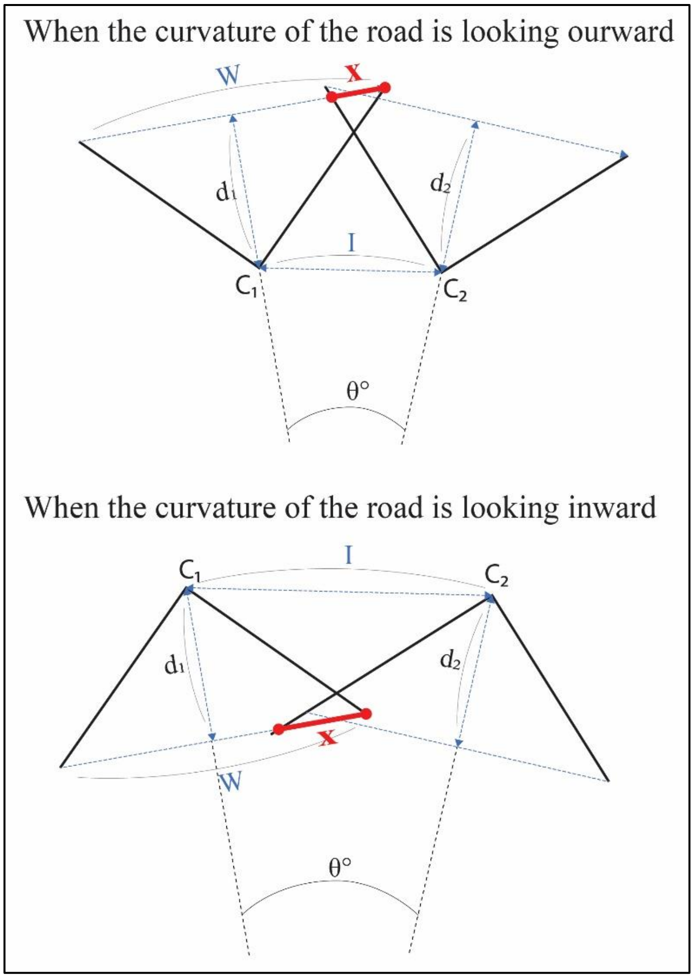


Figure 6. Method for estimating the proportion of the image overlapped (or gapped) with the next image

4.3.3.3 Cropping out the overlapped portion or append the gap by downloading additional images with rotated camera heading

Based on the estimated x/W, the images are cropped (or appended) to reduce the issue of double-counting (or accidentally omitting) items on street view images. If cropping is applied, the cropped image replaces the original image. When an additional image with a rotated camera heading is downloaded, the original image and the rotated image are likely to have overlapping portions. To prevent double counting issue from these overlaps, the original image and the rotated image were stitched into one image and used the stitched image to supplement the original image. Stitching was done using OpenCV 3.3.0 package in Python. Note that, among the two consecutive images, the cropping and appending is always applied to the image on the left (i.e., C_1 in Figure 6), as x/W is calculated to represent the proportion of C_1 image that needs to be cropped (i.e., when $x > 0$) or appended (i.e., when $x < 0$).

4.3.3 Computer vision

This chapter mainly uses Mask R-CNN[3], a semantic segmentation architecture by He et al. (2017), for its ability to (1) detect various objects within an image, (2) tell apart one object from others even among objects of the same kind (needed to count the occurrence of, e.g., streetlights), and (3) provide pixel-level masks for each object.

Although deep learning techniques applied on image data have shown remarkable success in extracting valuable information, their effectiveness is usually dependent on the size of the training dataset (Jean et al., 2016). For example, He et al. (2017) trained Mask R-CNN using the Microsoft COCO (Common Objects in Context), a large dataset for training computer vision models that contains 165,482 train, 81,208 validation, and 81,434 test images (Lin et al., 2014), to demonstrate the performance of their architecture and to distribute pre-trained weights. The challenge of this chapter was that there is no existing labeled dataset that contained all items on the MAPS-mini and that creating as large a dataset as the Microsoft COCO is infeasible. Although past implementations of

---

[3] https://github.com/matterport/Mask_RCNN

computer vision architectures indicate that they can be trained to detect items on the MAPS-mini, the lack of a large training dataset makes the application of these techniques challenging.

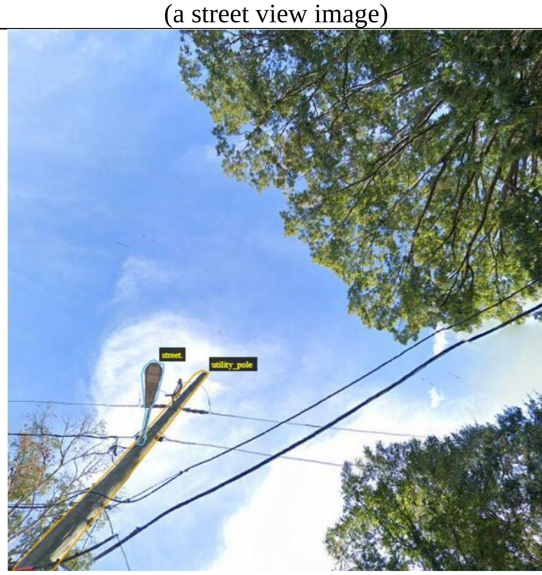
To overcome this challenge, this chapter uses a technique called ‘transfer learning.’ Transfer learning loads existing weights that are trained on a large dataset, such as Microsoft COCO or ImageNet, on the model architecture. Next, the layers that are responsible for the prediction of, in the case of Mask-RCNN as an example, class categories, bounding boxes, and masks are replaced with new layers with untrained weights. In the training process through transfer learning, only these new layers are trained; all other layers which are loaded with weights trained on a large dataset are frozen in the training process to take advantage of their ability for image feature extraction. This training method significantly reduces the number of parameters that need to be trained and allows users to ‘borrow’ the performance of a model trained on a large dataset and repurpose the model for a different task with a much smaller training dataset.

This chapter trains three separate models for Segment part images, Crossing part images, and the vertical view of Segment part images, for the ease of labeling the training dataset. Example images of the labeled training images are shown in Figure 7. The number of images labeled for the training and validation of computer vision models is around 2,500 with 20 classes for the Segment part of the MAPS-mini, about 500 images with 9 classes for the Crossing part, and 850 images with three classes for vertical view of Segment part. Street view images from random locations in Atlanta were sampled to create the training dataset. The random locations did not include street segments used for the validation. Because some items on the audit tools are rarely observed in the training dataset, a few non-street view, generic images on building maintenance condition and trip hazards were acquired through the internet search into the training dataset. A few hyperparameters were modified from the default setting provided by He et al. (2017), and image augmentation was heavily used to compensate for the small training dataset. The hyperparameters used in transfer learning of the Segment part are shown in Table 11 in Appendices.

(a) Segment – horizontal view
(a street view image)

(b) Segment – horizontal view
(a generic image from Google search engine)

(c) Segment – vertical view
(a street view image)



(d) Crossing
(a street view image)

Figure 7. Example images of the labeled training data

In addition to Mask R-CNN, this chapter also uses PSPNet to answer, “What percentage of the length of the sidewalk/walkway is covered by trees, awnings, or other overhead coverage?” As demonstrated in Chapter 3, PSPNet that is pre-trained on ADE20K can detect trees at pixel level (see Figure 3). PSPNet was applied to the images that are looking vertically up to estimate the proportion of the image that is covered by trees.

4.3.4 Geographic Information Systems

Four items in the MAPS-mini are audited using GIS and publicly available datasets. The GIS shapefiles of zoning designation, public parks, public transit stops, and bike paths were downloaded from Atlanta Regional Commission's data portal. The main 'type of land use' is determined by intersecting 15-meter buffer of street segments with the zoning shapefile. If more than 50% of the intersected area is commercial uses, the segment is classified as commercial. If otherwise, it is classified as residential. Public parks and transit stops are considered to be adjacent to a given segment if they fall into 15-meter buffer of the segment. For bike paths, midpoints of segments are generated, and 5-meter buffers are created using the midpoint. If there is a bike path intersecting with the buffer, the type of the intersecting bike path (e.g., painted line or physical barrier between bike land and road) is assigned to the segment.

4.3.5 Validation

This chapter uses stratified random sampling to select 100 street segments in Atlanta, GA, for validation. The four strata are defined using Walk Score and poverty rate at the Census Tract level. The Walk Score is used to reflect the fact that in U.S. cities, typical streetscapes in urban and suburban settings (e.g., as measured by Walk Score) tend to be distinct. The quality of microscale streetscapes can also vary depending on poverty rate even among areas with similar macroscale walkability (Bereitschaft, 2017; Neckerman et al., 2009). Using the median values of Walk Score and poverty rate, the strata are defined as ‘High Walk Score – High poverty rate,’ ‘High Walk Score – Low poverty rate,’ ‘Low Walk Score – High poverty rate,’ and ‘Low Walk Score – Low poverty rate.’ Proportional to the relative counts of Census Tracts that fall into each stratum, the total of 100 street segments are randomly selected. Figure 8 shows the location of four strata used for sampling streets as well as the location of the selected streets. Before the actual audit, two test auditors were recruited to conduct a pilot test of MAPS-mini audit on 15 street segments. Feedback from the pilot test auditors contributed to the development of an audit guideline, which was used to train the primary auditor. This auditor used Google Maps to audit 100 selected street segments using all functionalities available, including zooming and panning.

The images for the same streets were processed using the computer vision models and GIS. Before calculating how well the results from virtual audit and automated audit agree with each other, the information from computer vision models and GIS needed to be aggregated to street-level because, as illustrated in Figure 4, there are multiple images for each street. To do so, the count of the item of interest is summed up for each street segment and converted into respective categorical answers. For example, “How many public transit stops are present?” is answered by summing up the number of transit stops for a given street and converting the number into “0”, “1”, “2 or more” categories. For sidewalks and buffers, a given street segment is considered to have a sidewalk when sidewalks are detected in at least two consecutive images on the segment. Buffers detected in images without sidewalks are not counted. Because the MAPS-mini is about the pedestrian experience, only the trip hazards that are detected on sidewalks are counted.

The reliability of automated audit is assessed using percent agreement and Cohen’s Kappa (McHugh, 2012). Statistical analysis is conducted in R 4.0.2. Cohen’s Kappa is calculated using psych 2.0.8 package. Streetlight is an exception, as this chapter modified the question of “Are streetlights installed?” to have a numeric answer rather than a categorical one. This chapter conducted correlation analysis to assess the agreement for this item. Similarly, the total point has a numeric value and is assessed using correlation analysis.

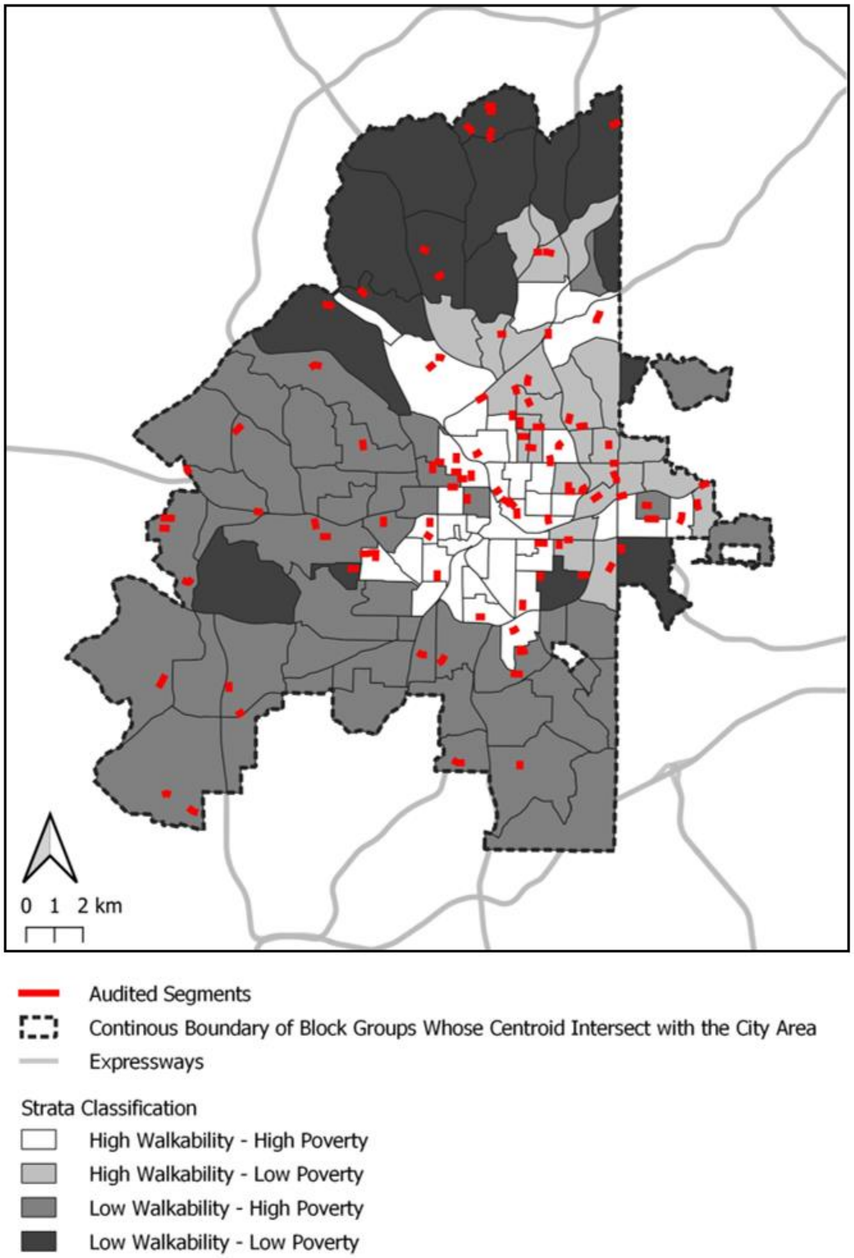


Figure 8. Four strata by Walk Score and poverty rate overlayed with the selected street segments

4.4 Results

Of the 100 street segments, four street segments had severe mismatches in road configuration between TIGER shapefile and Google or were missing GSV images. These segments were excluded. The average length of the 96 streets is 125.1 meters with standard deviation of 54.2 meters. The street length ranged between 51.3 and 299.4 meters. The street view image used in virtual and automated audits are temporally well-aligned: The virtual audit used images from 2016 – 2021 with the mean of 2019.28 and standard deviation of 0.790. The images used in automated audit are taken between 2015 to 2020 with a mean of 2019.12 and standard deviation of 0.847.

The total number of images used for the validation is 3,386, with about 35 images per street segment. Among the 3,386 images, 2,998 were for the Segment part, of which 996 are images looking upward, and 388 images for the Crossing part. Figure 9 shows example images of the prediction results that the computer vision model generated. The mean average precision (mAP), a commonly used metric for the performance of computer vision models, with the intersection over the union (IoU) threshold of 0.5 is 0.582 for the Segment part, 0.668 for the Crossing part, and 0.815 for the vertical view of the Segment part.

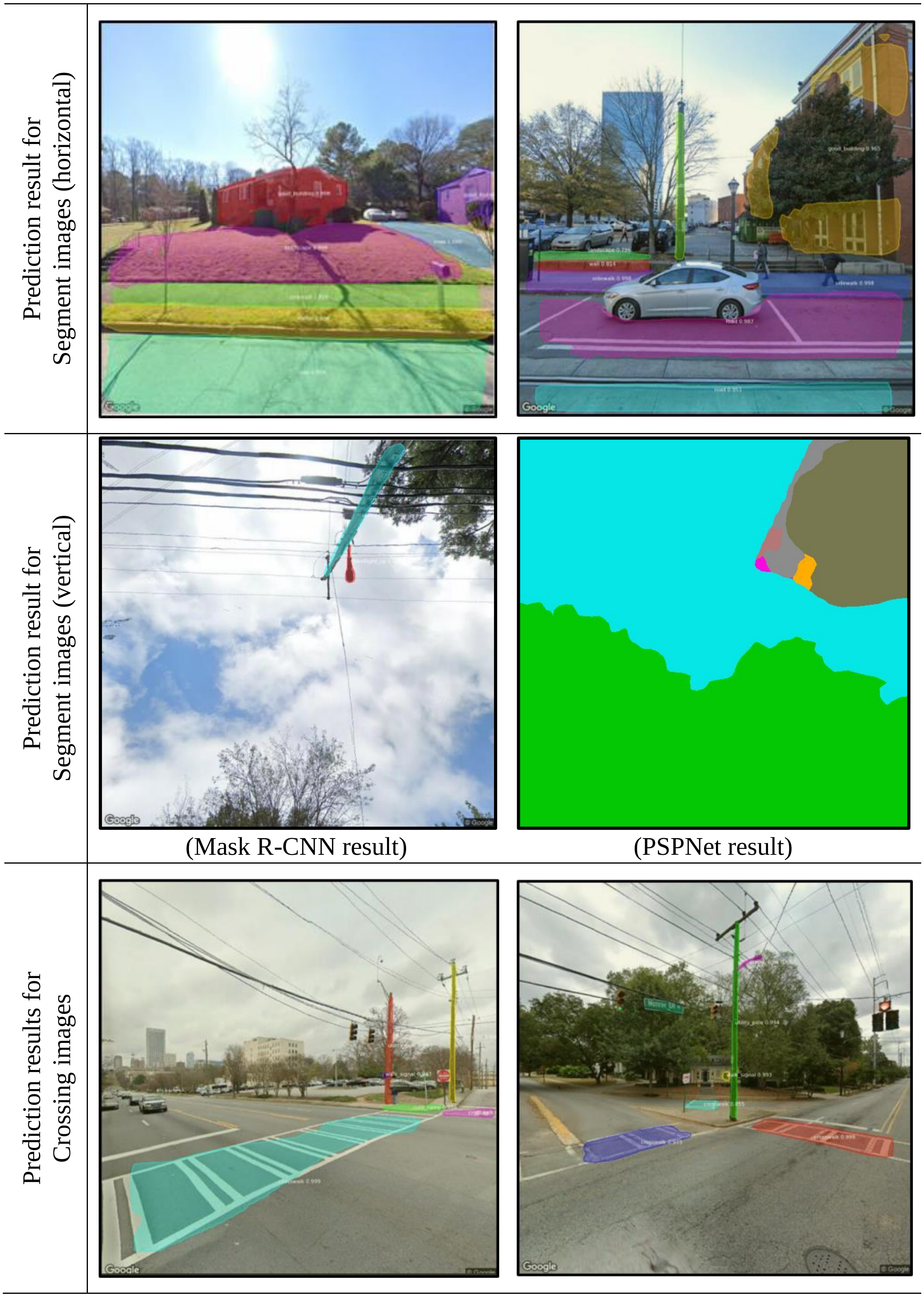


Figure 9. Examples of prediction results from the computer vision models

Of 2,002 images (i.e., 2998 – 996) that were considered for cropping or appending, 945 image locations had overlaps with the next images, and 975 image locations had gaps before the next images. For the overlapped images, about 28% of the left image needed to be cropped on average to eliminate the overlaps. For the images with gaps, about 47% worth of the left image needed to be appended to fill the gap with the next image. Figure 10 shows examples of the images with gaps and overlaps and the result of the cropping and appending process. In Figure 10(a), two consecutive images are taken at a close distance to each other, resulting in a substantial part of the street being captured in both of the images. For these two images, the method described above estimated that about 55% of the image of the left needs to be cropped to reduce the duplicate information. The red dotted box illustrates the portion of the image that needs to be cropped, and the green box illustrates the portion of the image that was used in the actual analysis. Figure 10(b) shows the case of a gap, where two consecutive images are taken too far from each other, resulting in a portion of the street not being captured in either of the images. The adjusted image displayed in the middle shows that there is a utility pole that would have been not detected if the original images were used for the analysis. The final analysis used the adjusted image. Note that 138 of these images were excluded from appending because the panorama-ID of GSV images were unstable and returned images from incorrect years, resulting in failures of the stitching algorithm.

(a) Example of two consecutive images with an overlap

(b) Example of consecutive images with a gap between the two

Figure 10. Example cases of images with overlaps/gaps.

Table 5 shows the observed agreement and Kappa statistics with 95% confidence interval between virtual and automated audits for categorical items. Observed agreements of the Segment items are high (> .80) except for the presence of ill-maintained buildings, indicating high reliability between most of the Segment items between virtual and automated audits. Kappa coefficients of the corresponding items are generally lower than observed agreement, especially for items that are less commonly observed (i.e., the presence of graffiti, seating, and trip hazards). The item with the lowest level of agreement is the presence of ill-maintained building (observed agreement = 0.646, κ = 0.171). The items audited using GIS consistently show high levels of agreement, even with the items that are rarely observed (observed agreement = 0.864 – 0.990, κ = 0.620 – 0.852). The correlation coefficient for the number of streetlights between the virtual and the automated audit shows a moderate correlation ($r(94) = 0.573$, $p < 0.001$).

Similarly, the three Crossing part items show moderate to high observed agreements. The presence of walk signal shows a high observed agreement with Kappa coefficient between 0.777 and 0.897. The level of agreement for the presence of crosswalks and curb ramps indicates that these items are less reliably audited (observed agreement = 0.719 – 0.760, κ = 0.401 – 0.484). Finally, the correlation coefficient for the total points between virtual audit and automated audit shows a high correlation ($r(94)$ = 0.752, $p < 0.001$).

Table 5. Levels of agreement between virtual and automated audit

Class	Item (Presence of..)	Observed agreement or correlation	Kappa	Kappa CI
Segment	Buffer	0.833	0.658	0.509 – 0.807
	No graffiti	0.906	0.351	0.024 – 0.679
	Seating	0.958	0.314	-0.178 – 0.806
	Sidewalk	0.896	0.717	0.556 – 0.878
	No trip Hazard	0.823	0.379	0.141 – 0.617
	No ill-maintained Building	0.646	0.171	-0.020 – 0.362
	Shade from overhead tree	0.688	0.357	0.196 - 0.518
	Streetlight	0.573[1]	-	-
	Bike Path	0.990[2]	0.852	0.566 – 1.000
	Public Park	0.947[2]	0.764	0.574 – 0.953
	Contains Commercial Uses	0.875[2]	0.728	0.584 – 0.872
	Transit Stop	0.864[2]	0.620	0.489 – 0.752
Crossing 1	Walk Signal	0.958	0.777	0.567 – 0.987
	Crosswalk	0.719	0.401	0.210 – 0.582
	Curb Ramp	0.719	0.435	0.255 – 0.615
Crossing 2	Walk Signal	0.979	0.897	0.757 – 1.000
	Crosswalk	0.740	0.484	0.312 – 0.655
	Curb Ramp	0.760	0.481	0.300 – 0.662
	Total Point	0.752[1]	-	-

[1]Agreement for streetlight is measured using Pearson's correlation.
[2]These items are audited using GIS as there commonly exists public GIS dataset for these items.

4.5 Discussion and conclusion

Compared to virtual audit, this chapter shows that many audit items in the MAPS-mini can be reliably audited using the combination of computer vision and street view images and GIS. Of 16 categorical items in Table 5, 11 items show high observed agreement (observed agreement > 0.80), and the rest are of moderate agreement (observed agreement > 0.60). Kappa coefficients tend to be lower than observed agreement, particularly for rare and/or qualitative items, such as the presence of ill-maintained buildings. The levels of agreement between virtual and automated audits are generally on par with the results of similar items from the past studies (e.g., Clarke et al., 2010; Griew et al., 2013). For some items, such as the presence of ill-maintained buildings and trip hazards, levels of agreement are higher than similar items reported in the literature (e.g., ‘any abandoned, burned out, or boarded up housing’ with $\kappa = 0.147$ and ‘street condition’ with $\kappa = 0.032$ in Clarke et al. (2010)). It is demonstrated that computer vision combined with street view images has the potential to offer a reasonably reliable and highly scalable alternative to virtual audit at a significantly lower cost. The fact that this chapter used small training datasets for the computer vision algorithms suggests that the performance of automated audit has the potential to be significantly improved with bigger training datasets.

While the overall level of agreement is acceptable, some items have relatively low levels of Kappa coefficients. One reason for some of the low level of Kappa coefficients is the rare occurrences of those items in both the training dataset for the computer vision algorithms and the audited segments. There are only four seating and six graffiti found in the audited segments. This rarity can result in a high chance agreement, which can decrease the magnitude of Kappa coefficient (Sim & Wright, 2005). For example, the chance agreement of seating is very high at 0.939. Plugging this value into the equation for the Kappa coefficient, which is (observed agreement – chance agreement) / (1 – chance agreement), it is apparent that this item requires an exceptionally high observed agreement to increase the Kappa coefficient. The rarity also means that these items appeared in training data less frequently, and the computer vision algorithm did not have enough to learn from.

Another reason for some of the low Kappa coefficient is rooted in the subjective nature of some item definitions. For example, the maintenance quality of buildings is intrinsically a continuous range with no objective cutoff line that separates good and bad maintenance quality. Labeling houses into either ill-maintained or normal buildings requires drawing an arbitrary cutoff line on the continuous scale. Despite the efforts to make the cutoff line as objective as possible by providing detailed examples in the audit guideline, this subjectivity is difficult to eliminate. This issue of subjectivity applies to graffiti (i.e., how much pleasing should it be aesthetically in order to qualify as a mural?) and trip hazard on sidewalks (i.e., how serious a damage should it be in order to be considered as a trip hazard?). Furthermore, trip hazards are not a single object but rather ill-maintained parts of sidewalks that can take a variety of appearances ranging from severely cracked sidewalks to overgrown grass that is tall enough to impede walking. This inconsistency in appearance can make detection more challenging. Walk signal is an example that illustrates the challenge posed by subjectivity and inconsistency in appearance. Because walk signals are objectively identifiable with nearly identical appearances in most cases and are clearly distinguishable from other objects that are not walk signals, distinguishing them is less ambiguous both for the human auditor and the computer vision algorithms, resulting in the high reliability score in Table 5.

Note that these issues of rarity and subjectivity are not intrinsic to computer vision algorithms and/or GSV images. The rarity of some items, such as graffiti or seating, is likely to be characteristics of the streetscapes specific to Atlanta and may not be attributable to GSV images. Also, past studies showed that items requiring subjective assessment tend to result in low levels of reliability even among human raters (Clarke et al., 2010; Clifton et al., 2007).

This chapter identified a few technical challenges associated with using computer vision algorithms and GSV images for streetscape audits. First, the maximum resolution of images that Google Street View API provides is 640 by 640 pixels. While this is a sufficient resolution when objects of interest are located close to the road from which photos were taken, objects that are located far from the road or small in size can appear blurry in GSV images. This blur can obscure the shape and texture of objects and potentially have negative impacts on the prediction accuracy of computer vision models. Examples of items that can be influenced by this issue include walk signals, trip hazards, seating, curb ramp, and single-family houses with a large front yard. Second, the street view images are limited to public roads and are fixed at the centerline of the road. As mentioned in previous studies, items that are on or closer to the ground (e.g., trip hazards, seating) can be challenging to audit using street view images because they can be hidden behind other objects such as parked cars and tree trunks (Rundle et al., 2011). Third, there are many locations where the distance between two consecutive images is long (or short) enough to create a significant gap (or an overlap) between the two images, assuming that the FOV is set to 90 degrees. The gaps and overlaps are frequently observed not only on curvy segments but also on completely straight segments. This issue requires careful processing before performing automated environmental audits using GSV images.

Furthermore, the author found no obvious ways to predict a priori whether a segment has gaps or overlaps, which requires that the image collection and processing should be at least a two-stage process. In the first stage, image collection and processing through computer vision algorithms are done without any information about the presence of gaps and overlaps. The result of this stage generates the estimation of gaps and overlaps. Based on the estimation, the overlapped portions of the collected images are cropped, and additional images covering the gaps are collected and appended to the images from the first stage. These ‘corrected’ images are then processed using computer vision algorithms to generate the final audit result. Furthermore, even after the gaps and overlaps are corrected properly, the objects located on the gapped portion of a segment can only be captured by rotating the heading of the nearest image location. This often means a longer distance between objects and the camera. The distance combined with the low resolution can reduce the information with which computer vision algorithms can perform detection.

The limitations of the study can provide insights on how future studies can improve upon this chapter. Due to the limitations in the available resources, increasing the size of the training data was infeasible. As one of the most important factors that gave a boost to deep learning-based computer vision algorithms is the appearance of large labeled datasets (Voulodimos et al., 2018), increasing training data to a sufficiently large size may provide substantial improvements to the effectiveness of automated audits, particularly for items that were rarely found in our data (i.e., graffiti and seating). Second, this chapter did not assess the reliability between in-person and automated audits. Past studies have demonstrated that virtual audit can reliably assess the streetscapes except for items that are temporally variable or small in size (Clarke et al., 2010; Rundle et al., 2011). Several studies have already used virtual audits as a sole source of information on the streetscapes without conducting in-person audits and linked the virtual audit result to outcome variable of their interest (e.g., Hanson et al., 2013; L. He et al., 2017; Mooney et al., 2016). The goal of this chapter is to develop an automated method that can replicate these virtual audits. Third, the labeling of the training dataset for the computer vision algorithm can be further refined. The maintenance quality of buildings, trip hazards, seating, and graffiti categories may benefit from a more detailed labeling strategy. For example, in addition to labeling an entire house as an ‘ill-maintained building’ or ‘well-maintained building,’ future studies may benefit from adding labels of individual components that collectively constitute the building maintenance quality, such as boarded windows or cracked outer structures, and factors that could lead to the binary classification of maintenance quality.

This chapter demonstrated the effectiveness of using GSV images and computer vision algorithms for automatically auditing the streetscapes. This chapter identified several challenges specific to Atlanta streetscapes and to GSV images, respectively, which could have been contributing factors to the lower levels of agreement observed for some items. However, despite these challenges, the results of this chapter demonstrate that computer vision algorithms and GSV images can provide a reliable method for automatically auditing streetscapes. If the challenges identified in this chapter are mitigated and overcome in future studies, the method is expected to offer a cost-effective method for conducting a streetscape audit using a scientifically validated audit tool (i.e., the MAPS-mini) at regional or even at national scale, a task that has been prohibitively expensive if using conventional methods.

CHAPTER 5. UNPACKING THE ASSOCIATIONS OF WALKABILITY FACTORS OF DIFFERENT SCALES WITH DISADVANTAGED GROUPS' WALKING BEHAVIOR

5.1 Introduction

The literature on the association between the built environment and walking behavior have now established that certain characteristics of the built environment can enable or constrain the opportunities for walking. Chapter 3 examined the association between macro- and mesoscale factors of walkability with walking mode choice, and this chapter shifts its focus to how *microscale* factors are connected with walking mode choice.

Microscale factors of walkability are the most fine-grained design details that are attached to the three-dimensional structure of streetscapes. They also include the quality of the design details. Microscale factors, such as buffers separating sidewalks from the road, ample streetlights, street furniture, and the maintenance quality of sidewalks, can improve the perceived attractiveness of the street (Adkins et al., 2012; Borst et al., 2008) and can facilitate walking and active transport (M. Alfonzo et al., 2008; Handy et al., 2002; Sallis et al., 2015) by providing safety, comfort, and pleasurability to pedestrians. While past studies tended to put greater emphases on macroscale factors than on microscale factors due to, in part, the difficulty of measuring microscale factors, recent findings are providing support for the importance of microscale factors even after controlling for macroscale factors (Adkins et al., 2012; M. Alfonzo et al., 2008; Cain et al., 2014; Ewing & Clemente, 2013).

Microscale factors of walkability can be particularly important for disadvantaged populations' walking behavior, as disadvantaged populations tend to be less responsive to macroscale factors or, in some cases, have opposite responses to the expected effects of macroscale factors (Adkins et al., 2017; Forsyth et al., 2009; Frank et al., 2008; Lovasi, Hutson, et al., 2009; Manaugh & El-Geneidy, 2011). Macroscale factors (e.g., residential density, land use mix, intersection density, or Walk Score) reported weaker association with walking among nonwhite and low-income youth (Kerr et al., 2007) and among black male (Frank et al., 2004) or insignificant association with walking among low-income populations with no cars (Manaugh & El-Geneidy, 2011). The insensitivity of disadvantaged populations to macroscale factors of walkability can limit the applicability of the findings from past studies that primarily use macroscale factors.

The insensitivity to macroscale factors among the disadvantaged populations can arise through two main pathways (Adkins et al., 2017). First, disadvantaged populations can walk more than expected in an environment with unsupportive macroscale walkability factors. This can happen due to limited choices on travel mode (e.g., having to walk to places due to limited vehicle access). Second, they can walk less than expected in an environment with supportive macroscale walkability factors due to barriers in the social and physical environment such as concerns on safety from crime and traffic, and aesthetic problems, or other social influences that override the impact of macroscale factors (Adkins et al., 2017; Bennett et al., 2007; Lovasi, Neckerman, et al., 2009). Microscale factors of walkability can be relevant to disadvantaged populations' insensitivity to macroscale factors through the second pathway. The concerns for safety, comfort, and aesthetic or pleasurability can be closely linked with microscale factors (Alfonzo, 2005), and studies reported that disadvantaged neighborhoods often have unfavorable microscale walkability factors compared to more advantaged populations (Bereitschaft, 2017; Neckerman et al., 2009; Sallis et al., 2011).

What is less understood is which of the two pathways for the weaker relationship is more significant. Note that one of the two pathways pertains to the characteristics of individuals and households that make trips (i.e., the limitations on mode choice options) while the other pathway relates to the environment in which they make trips (i.e., the physical environment that lacks safety, comfort, and pleasurable qualities). While planners cannot directly intervene on the limited mode choice status of individuals and households, interventions on the physical environment for better safety, comfort, and pleasurabilty fall into the jurisdiction of urban planners and designers. Therefore, a more detailed understanding of how the weaker relationship arises is needed for more effective urban planning and design interventions.

Many past studies used socio-demographic variables that can encapsulate both pathways to characterize disadvantaged groups, such as race and education attainment. Some studies that did consider specific pathways for the weaker relationship often included one pathway or the other, but not both pathways simultaneously in one research framework (e.g., Forsyth et al., 2009). Due to these limitations, past studies were insufficient in providing a detailed understanding of exact reasons for the insensitivity of disadvantaged populations and were limited in providing guidance to planners who wish to encourage walking among disadvantaged populations.

This chapter attempts to answer the following research questions: "How are microscale factors of walkability associated with macroscale factors?" As discussed in Chapter 3, if microscale factors are highly correlated with macroscale factors, they may not provide additional useful information. The second research question is "which of the two hypothetical pathways more prominent?" This chapter expands the third chapter by incorporating how the weaker relationship between macroscale factors of walkability and disadvantaged populations' walking behavior arises by using variables that correspond with each of the two pathways. Because there are little empirical knowledge and theoretical hypotheses in the literature on the effect of mesoscale factors of walkability on disadvantaged populations, this chapter focuses on macroscale and microscale factors of walkability.

After an overview of the literature on the relationship between disadvantaged population and walkable built environment, this chapter examines the descriptive statistics for people with limited mode choice status and those without the limitation. Then, correlations between individual microscale factors of walkability and macroscale index are examined. In the main regression modeling, this chapter empirically validates whether automatically audited microscale factors contribute to walking mode choice models by entering microscale factors into a series of binary logistic regression models on walking mode choice. Finally, it examines how and whether the effects of macroscale index on walking mode choice are moderated by limited mode choice status and microscale index.

5.2 Literature Review

There are two opposing hypotheses in the literature on how disadvantaged populations respond to the walking-related effects of the built environment. One hypothesis is that disadvantaged populations are more responsive to their residential environment than their counterparts because they do not have the means to leave their residential environment and are more exposed to their residential environment (Ivory et al., 2015; Lovasi, Hutson, et al., 2009, p. 279). The opposing hypothesis states that disadvantaged populations are less responsive to the environment in which they travel. (Lovasi, Neckerman, et al., 2009). The evidence in the literature seems to offer more support to the second hypothesis that disadvantaged groups are less responsive to the built environment (Adkins et al., 2017; Lovasi, Hutson, et al., 2009). A review paper by Adkins et al. (2017) succinctly summarizes the relevant literature published between 2004 and 2015. Of the 17 studies reviewed, 14 studies found weaker effects of the built environment for disadvantaged groups. The difference in the effect between disadvantaged and advantaged groups was notable: "… the average built environment effect on transport walking for advantaged groups is 2.6 times stronger than it is for disadvantaged groups" (Adkins et al., 2017, p. 307). For physical activity and leisure walking, the stronger relationship between the built environment and walking for advantaged groups was also observed but was less prominent than it was for transport walking (Adkins et al., 2017).

The weak relationship between walking and the built environment can be caused by either or a combination of the following two hypothetical pathways. First, disadvantaged populations can walk more than expected despite an unsupportive environment (i.e., unsupportive macroscale walkability factors) if their mode choice options are limited (Adkins et al., 2017; M. A. Alfonzo, 2005). Because they need to travel to meet their daily needs (e.g., school or work, groceries, and banks), those with limited mode choices may be unable to avoid the unsupportive built environment by driving and may be more likely to walk more frequently or further than expected compared to those in the similarly unsupportive built environment but have access to vehicles (Lovasi, Neckerman, et al., 2009). Similarly, Alfonzo (2005) argues that if the choice is limited to walking, the hierarchy of walking needs play little role, and so do the corresponding characteristics of the built environment. The issue of choice may be particularly important for economically disadvantaged populations (M. A. Alfonzo, 2005), as the limited choice is often linked with car ownership or vehicle availability.

The second hypothetical pathway for the weaker relationship is that disadvantaged populations can walk less than expected in a supportive built environment (i.e., supportive macroscale walkability factors) if there are barriers to walking, such as concerns about safety from crime and traffic or aesthetic disorderliness, which can override the positive effects of supportive macroscale walkability factors and discourage walking and physical activity (Adkins et al., 2017; Bennett et al., 2007; Day, 2006; Harrison et al., 2007; Hooker et al., 2005). Even in areas with well-connected street patterns with a density of diverse destinations, concerns for safety from crime, lack of comfort from traffic, and aesthetic disorderliness may suppress the supportiveness of macroscale factors.

Disadvantaged populations are more likely to have restricted vehicle access and be limited in mode choice options. They are also likely to be disproportionately exposed to unfavorable environments in terms of safety, comfort, and pleasurability. On the exposure to unfavorable environment, Sallis et al. (2011) found that “residents from high-income neighborhoods reported more favorable esthetics, pedestrian/biking facilities, safety from traffic, safety from crime, and access to recreation facilities than residents of low-income areas (all p’s < 0.001)” (p. 1274). Similarly, Neckerman et al. (2009) found that, after controlling for macroscale walkability, poor neighborhoods have “significantly fewer street trees, landmarked buildings, clean streets, and sidewalk cafes, and higher rates of felony complaints, narcotics arrests, and vehicular crashes” (p. S264). In a comparison between streetscapes with high and low social vulnerability but similar Walk Score, Bereitschaft (2017) reported that “streetscapes in neighborhoods with high social vulnerability exhibited less contiguous street walls, fewer windows and less transparent storefronts, less well maintained infrastructure, fewer street cafés, and overall less complexity than those in neighborhoods with low social vulnerability” (1). Specifically on safety from traffic, several studies reported that disadvantaged populations disproportionately suffer from exposure to unsafe traffic conditions or heavy traffic (Chichester et al., 1998; Huston et al., 2003; King & Palmisano, 1992), which can discourage walking.

It is worth reiterating that most of the past empirical studies used macroscale factors to examine how disadvantaged groups respond to the walking-related built environment. Only three out of 17 studies in a review by Adkins et al. (2017) included street-level variables, and the rest used some combination of intersection density, residential density, retail floor area ratio, land use mix, the density of business and services, and Walk Score (Adkins et al., 2017). The measures commonly used in conventional walkability indices are "… developed and tested largely in the context of relatively advantaged communities…" (Adkins et al., 2017, p. 299) and are missing some important walking-related factors, such as concerns for safety and disorderliness (Adkins et al., 2017).

As mentioned earlier, most past studies characterized disadvantaged populations using socio-demographic/economic characteristics, such as age, income, ethnicity, sex, car ownership, or education attainment (e.g., Frank et al., 2008), and few incorporated variables that can directly test which of the two pathways is responsible for the weaker relationship. Because these socio-demographic/economic characteristics are outside of planners' jurisdiction, past studies are limited in offering intervention strategies for disadvantaged populations from planning perspectives.

In summary, although disadvantaged populations tend to respond less or inversely than expected to walkable macroscale factors, less is understood about which of the two hypothetical pathways for the weaker relationship is more prominent. Past studies were limited by the use of generic socio-demographic/economic characteristics for characterizing disadvantaged populations. This chapter contributes to the literature by tapping into each of the two pathways for the weaker relationship between disadvantaged populations and walkable macroscale factors. This chapter does so by including interaction terms that consist of specific variables for each of the two pathways, namely the limited mode choice options and microscale factors.

5.3 Data and Analytical Methods

5.3.1 National Household Travel Survey data

The National Household Travel Survey data (NHTS) was collected and processed through the same method described in Chapter three. One additional variable was calculated for this chapter: limited mode choice status. It is defined as a binary categorical variable where it equals 1 if the person belongs to a household that has less than one car per adult or if the person walks to reduce the financial burden of travel (i.e., answering “strongly agree” to a question “do you walk to save money?”). It is given 0 otherwise.

5.3.2 Macroscale Factors and Mesoscale Factors

The macroscale factors and mesoscale factors of walkability were collected and processed through the same method described in Chapter 3. An important difference is that instead of using the individual macro and mesoscale factors, this chapter combines the macroscale factors and mesoscale factors into macroscale index and mesoscale index, respectively, by converted them into z-scores and summing them up (Frank et al., 2005). This was to limit the number of independent variables as well as to make the implementation and interpretation of the interaction terms easier. Before converting to z-score, distance to transit station variable was inverted by subtracting each value of the variable from its own maximum value to reflect the fact that proximity to transit represents higher walkability. This inverted variable is used for the analysis of this chapter (heretofore, it is called proximity to transit instead of distance to transit). Also note that ‘sidewalk-to-street proportion’ of mesoscale factors was excluded from this calculation because it was not statistically significantly associated with walking mode choice in the previous chapter and there is no prior research that confirmed its effectiveness.

5.3.3 Microscale Factors

This chapter adds microscale factors to the data used in Chapter 3. To apply the method developed in Chapter 4, a buffer was drawn around the origin location of each trip to select street segments relevant to the mode choice of the trip. Then, 150-meter was used as the buffer distance. Street segments are excluded if less than one-tenth of the street segment intersects with the buffer.

For the selected street segments, automated audit was conducted using the method developed in Chapter 4. Note that each item in the MAPS-mini assigns either 1 or 2 to an item if the item is found on a street and 0 if not. If the item measures an unfavorable feature for pedestrians (e.g., the existence of trip hazard), a street is given the value of 1 if the item is not found on the street. In addition to the 15 individual items of the MAPS-mini, a microscale index was calculated by summing up the score of each item.

Because there can be multiple streets within the buffer of a trip origin, item scores and microscale index for each trip origin are calculated by averaging the item and the index scores of street segments around the origin location. For example, if there are five street segments that fall into the buffer of a given trip origin location and if three of the streets had sidewalks detected, the sidewalk score for the trip origin location would be three divided by five. Similarly, if there are three street segments with microscale index of 10, 12, and 13, the microscale index for the location would be 11.667 (i.e., (10+12+13) / 3).

Note that when downloading street view images through the GSV API, it is not possible to specify which year of image the API would return, and the API can return different images for the same location if Google has updated images for that location. Because images for microscale factors were downloaded later than those for mesoscale factors, mesoscale factors were calculated with images taken earlier on average than those for microscale factors. Most of the images used for mesoscale factors were taken between 2016 and 2018, while most of the images for microscale factors were taken between 2016 and 2020.

5.3.4 Analytical Methods

This chapter first examines the descriptive statistics of the data for subgroups characterized by the limited mode choice status, and t-test results are reported if the difference in mean between the subgroups is significant. To answer the first research question posed at the end of the introduction of this chapter, this chapter examines how macro, meso, and microscale indices are associated with one another using correlation analysis. A set of binary logistic regression models are then developed to examine whether the addition of microscale index to macro and mesoscale indices improves the ability of walking mode choice models in predicting the mode choice.

For the second research question, another set of binary logistic regression models with interaction terms were specified to test the two pathways for the weaker relationship between macroscale walkability factors and walking mode choice for disadvantaged populations. Note that for regression analysis, this chapter uses only microscale index rather than the 15 individual factors in the regression analysis because (1) the study has a limited sample size, (2) to make implementation and interpretation of interaction terms easier, and (3) the total point is reported to have a linear, positive, and significant association with active transport for all age groups (Sallis et al., 2015).

Specifically, the Base Model uses the mode choice (i.e., walk trip vs. non-walk trip) as the dependent variable and includes only the control variables as the independent variables. Model 1 through 6 add macro, meso, and microscale indices in various specifications to examine how each index contributes to walking mode choice models. Model 7 and 8 maintain all the variables in Model 6 (i.e., a model that includes control variables and all three indices) and add interaction terms that test each of the two pathways for the weaker relationship. Finally, Model 9 includes all variables (i.e., Model 6 and the two interaction terms in one model).

5.4 Results

5.4.1 Descriptive statistics by with and without limited mode choice status

Table 6 shows the mean and standard deviation of the variables for the two groups for all trips used in the regression analysis. People with limited mode choice status are, on average, about 12 years younger ($t(85.19) = 4.510$, $p < 0.001$) and earning about $40,000 less household income ($t(74.99) = 3.192$, $p < 0.002$). They also had higher proportion of non-white populations (53.5% non-white) compared to those without limitations (30.5% non-white), less than high school education (9.3% less than high school) compared to those without limitations (1.7% less than high school), and nondrivers (32.6% non-driver) compared to those without limitations (1.7% non-driver). Note that data presented in Table 6 used trip-level data (i.e., it can contain more than one trips per person) while the mean and the t-test results presented above in this paragraph used person-level data, which can result in slight differences.

T-tests for macro, meso, and microscale indices showed that the means of macroscale and microscale are not significantly different between the two groups (t(316) = -0.168, p = 0.866 for macroscale index; t(316) = 0.092, p = 0.927 for microscale index). However, the mean of mesoscale index for those with limited mode choices were lower than their counterpart (t(316) = 2.510, p < 0.013).

Table 6. Summary statistics of the variables (number of trips = 318)

Variable	**No limitations in mode choice**	**With limitations in mode choice**
Count	222	96
Age	Mean: 46.3 / SD: 17.0	Mean: 34.5 / SD: 13.6
Gender	Male: 134 (60.4%) Female: 88 (39.6%)	Male: 50 (52.1%) Female: 46 (47.9%)
Race	White: 165 (74.3%) Black: 45 (20.3%) Other race: 12 (5.4%)	White: 44 (45.8%) Black: 43 (44.8%) Other race: 9 (9.4%)
Education	Less than high school: 3 (1.3%) High school or higher: 219 (98.6%)	Less than high school: 8 (8.3%) High school or higher: 88 (91.7%)
Household income ($)	Mean: 113,771.8 / SD: 69,131.6	Mean: 56,900.5 / SD: 60,818.3
Driver status	Driver: 216 (97.3%) Non-driver: 6 (2.7%)	Driver: 62 (64.6%) Non-driver: 34 (35.4%)
Count walk trips in past 7 days	Mean: 9.1 / SD: 9.3	Mean: 11.4 / SD: 8.9
Travel distance (miles)	Mean: 0.444 / SD: 0.288	Mean: 0.430 / SD: 0.243
Macroscale index	Mean: -0.023 / SD: 3.9	Mean: 0.053 / SD: 3.291
Mesoscale index	Mean: 0.076 / SD: 0.794	Mean: -0.177 / SD: 0.892
Microscale index	Mean: 9.059 / SD: 1.474	Mean: 9.076 / SD: 1.370

5.4.2 Association between macroscale and microscale factors of walkability

Most of microscale factors showed significant levels of correlation with various macroscale factors (Table 7). While most microscale factors showed positive correlations with macroscale index, three microscale factors, including buffer, shade from overhead cover, and absence of graffiti, were negatively correlated with macroscale index. These negative correlations make sense: higher macroscale walkability factors suggest more compact use of land, which can limit spaces for buffer and tree planting on streets. Graffiti is often found on public properties, particularly in transportation facilities such as transit stations and shelters, or walls and other similar objects open to public view (Weisel, 2002). Central cities can have more locations that meet these conditions than, for example, residential subdivisions in the suburbs. Other items with positive correlations also align with the expectation. Items with the purpose of serving pedestrians are expected to be found frequently in areas with high macroscale walkability factors. It is also not surprising to see high correlations between crossing items (i.e., walk signal, crosswalk, and curb ramp) and compact development pattern that usually accompanies grid street network.

Table 7. Correlation between microscale factors and macroscale factors

	Macroscale index	Employment density	Land use diversity	Intersection density	Proximity to transit	Walk Score
Microscale index	0.469***	0.498***	0.234***	0.177**	0.356***	0.473***
Buffer	-0.209***	-0.247***	-0.240***	-0.33*	0.024	0.017
No graffiti	-0.331***	-0.313***	-0.142*	-0.098†	-0.274***	-0.398***
Seating	0.147**	0.176**	-0.033	0.114*	0.126*	0.162**
Sidewalk	0.276***	0.163**	0.105†	0.088	0.239***	0.426***
No trip hazard	0.096†	0.194**	0.017	0.147**	-0.006	0.002
No ill-maintained building	0.108†	0.032	0.000	0.067	0.115*	0.187***
Shade from overhead tree	-0.255***	-0.208***	-0.250***	-0.283***	-0.113*	-0.091
Streetlight	0.365***	0.369***	0.255***	0.150**	0.290***	0.290***
Bike path	0.229***	0.344***	0.062	0.158**	0.165**	0.122*
Public park	0.041	0.063	0.009	0.035	0.131*	-0.087
Contains commercial uses	0.599***	0.474***	0.419***	0.401***	0.392***	0.534***
Transit stop	0.228***	0.355***	0.226***	0.025	0.033	0.206***
Walk signal	0.516***	0.579***	0.252***	0.272***	0.355***	0.456***
Crosswalk	0.291***	0.284***	0.211***	0.179**	0.194***	0.211***
Curb ramp	0.302***	0.266***	0.066	0.228***	0.284***	0.274***

† Significant at the 10% level; * Significant at the 5% level; **Significant at the 1% level; *** Significant at < 1% level.

Table 8 shows the odds ratio and z-statistics of the first six binary logistic regression models examining the relative contribution of microscale index on walking mode choice. For ease of comparison, the odds ratio was calculated based on standardized coefficients. Average marginal effects of all regression results can be found in Table 12 and Table 13 in Appendices. Model 1 and 2 show that when macro and mesoscale indices are added to the Base Model (i.e., model with control variables), they each provide a significant contribution to the model fit (likelihood ratio test between Base Model and Model 1: $\chi^2(1) = 27.9$, $p < 0.001$; between Base Model and Model 2: $\chi^2(1) = 11.76$, $p < 0.001$), with coefficients that align with the expected direction of effect. Microscale index in Model 3 was marginally significant at $\alpha = 0.1$ (likelihood ratio test between Base Model and Model 3: $\chi^2(1) = 2.9$, $p = 0.091$). As shown in Model 4 and 5, microscale index becomes statistically insignificant when macroscale index enters the model, but it became statistically significant at $\alpha = 0.05$ when it is coupled with mesoscale index. When all three indices enter the model in Model 6, microscale index is no longer statistically significant. This result was expected given the significant correlation between macroscale and microscale indices. The result aligns with some of the past findings that emphasized the role of macroscale factors over microscale factors, as well as the hierarchical walking needs hypothesis, which poses that the need for accessibility is more fundamental than the needs for safety, comfort, and pleasurability (M. A. Alfonzo, 2005).

Table 8. Results of the logistic regression models with microscale index (dep.var = walking / non-walking in binary)

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Constant	1.802 (0.641)	0.345 (-1.130)	0.760 (-0.299)	1.857 (0.672)	0.367 (-1.03)	0.87 (-0.142)
Age	0.625* (-2.449)	0.536*** (-3.351)	0.547*** (-3.311)	0.618* (-2.487)	0.579** (-2.891)	0.638* (-2.284)
Gender – Male (Base: Female)	1.101 (0.278)	1.362 (0.934)	1.297 (0.800)	1.085 (0.234)	1.372 (0.943)	1.141 (0.372)
Race – Black (Base: White)	0.603 (-1.089)	0.963 (-0.091)	0.786 (-0.587)	0.619 (-1.03)	0.903 (-0.242)	0.714 (-0.717)
Race – Other races (Base: White)	1.121 (0.163)	4.305* (2.164)	3.132† (1.767)	1.100 (0.135)	4.312* (2.191)	1.435 (0.501)
Household income	1.298 (1.264)	1.574* (2.380)	1.634* (2.567)	1.296 (1.254)	1.515* (2.128)	1.211 (0.906)
Household vehicle count	0.508** (-2.84)	0.55** (-2.785)	0.508** (-3.006)	0.517** (-2.734)	0.522** (-2.918)	0.546* (-2.522)
Limited mode choice	0.261** (-2.946)	0.408* (-2.034)	0.311** (-2.649)	0.261** (-2.948)	0.406* (-2.016)	0.326* (-2.414)
Education – HS or higher (Base: Less than HS)	1.927 (0.764)	6.577* (2.093)	3.542 (1.439)	1.869 (0.728)	6.465* (2.013)	3.673 (1.414)
Driver status – Non-Driver (Base: Driver)	9.996** (3.039)	5.953* (2.526)	4.927* (2.311)	9.678** (3.006)	7.550** (2.681)	14.282** (3.242)
Number walking activities in the past 7 days	2.585*** (4.717)	2.419*** (4.621)	2.562*** (4.755)	2.592*** (4.737)	2.373*** (4.438)	2.434*** (4.499)
Trip distance	0.274*** (-6.429)	0.220*** (-7.579)	0.241*** (-7.357)	0.272*** (-6.446)	0.231*** (-7.256)	0.260*** (-6.443)
Macroscale index	2.947*** (4.633)			3.090*** (4.490)		2.869*** (4.087)
Mesoscale index		1.713** (3.097)			1.862*** (3.422)	1.681** (2.636)
Microscale index			1.317† (1.673)	0.903 (-0.526)	1.485* (2.263)	1.029 (0.140)
No. of observation	318	318	318	318	318	318
LL	-116.6	-125.6	-129.1	-116.5	-122.9	-112.9
Adj. McFadden's R^2	0.375	0.332	0.315	0.371	0.340	0.384
Bayesian Info. Criteria	308.2	326.1	333.2	313.7	326.5	312.2

- The regression results are in $\frac{Odds\ Ratio\ ***}{(z-statistic)}$ format, where the Odds Ratio is the exponent of the standardized coefficient from the logistic regression.
† Significant at the 10% level; * Significant at the 5% level; **Significant at the 1% level; *** Significant at < 1% level.

5.4.3 The two pathways for the weaker relationship between disadvantaged and the built environment

The results of binary regression models with interaction terms are presented in Table 9. Model 7 shows the result from the first interaction term, which consists of macroscale index multiplied by the limited mode choice status. The interaction term is statistically significant, adding a meaningful improvement to the model fit over Model 6 (Likelihood test between Model 6 and 7: $\chi^2(1)= 15.8$, $p<0.001$). When the second interaction term consisting of macroscale index and microscale index is used, it also offers a statistically significant improvement to the model fit over Model 6 (Likelihood test between Model 6 and 8: $\chi^2(1)= 4.8$, $p=0.029$), but the magnitude of improvement is smaller than the first interaction term consisting of macroscale index and limited mode choice status.

Finally, when all variables and interaction terms enter in Model 9, both interaction terms are significant, resulting in the best model fit (Likelihood test between Model 7 and 9: $\chi^2(1) = 13.1$, $p<0.001$). These results suggest that the two hypothetical pathways for the weaker relationship between disadvantaged populations and macroscale walkability factors are both in effect and that the limited mode choice status may be the more prominent pathway between the two. Yet, the concern for safety, comfort, and pleasurability as measured by microscale index plays a sizable role, and including the two pathways together in one model offers the best adjusted R^2 and BIC statistics.

Although it is not a part of the two hypothetical pathways, the possibility of interaction between microscale index and the limited mode choice status, as well as the possibility of three-way interaction, was tested. As they were both statistically insignificant, these results are not displayed in Table 9.

Table 9. Results of the logistic regression models with interaction terms (dep.var = walking / non-walking in binary)

	Model 7	**Model 8**	**Model 9**
Constant	1.057 (0.055)	0.909 (-0.099)	0.855 (-0.167)
Age	0.694† (-1.792)	0.606* (-2.483)	0.636* (-2.117)
Gender – Male (Base: Female)	1.072 (0.194)	1.270 (0.657)	1.253 (0.601)
Race – Black (Base: White)	0.596 (-1.040)	0.520 (-1.317)	0.292* (-2.226)
Race – Other races (Base: White)	1.389 (0.45)	1.199 (0.249)	0.981 (-0.025)
Household income	1.208 (0.876)	1.205 (0.855)	1.167 (0.685)
Household vehicle count	0.514** (-2.696)	0.515** (-2.648)	0.474** (-2.813)
Limited mode choice	0.296** (-2.587)	0.335* (-2.301)	0.262** (-2.711)
Education – HS or higher (Base: Less than HS)	3.830 (1.416)	3.312 (1.351)	5.106† (1.857)
Driver status – Non-Driver (Base: Driver)	10.321** (2.953)	12.408** (3.107)	10.141** (2.908)
Number walking activities in the past 7 days	2.325*** (4.342)	2.483*** (4.492)	2.346*** (4.184)
Trip distance	0.246*** (-6.252)	0.255*** (-6.472)	0.225*** (-6.321)
Macroscale index	5.492*** (4.947)	3.141*** (4.161)	8.810*** (5.341)
Mesoscale index	1.896** (3.024)	1.603* (2.356)	1.836** (2.702)
Microscale index	1.022 (0.100)	1.225 (0.899)	1.495 (1.564)
Interaction: Macro index * Limited mode choice	0.166*** (-3.872)		0.070*** (-4.441)
Interaction: Macro index * Micro index		1.566* (2.117)	2.523*** (3.355)
No. of observation	318	318	318
LL	-105.006	-110.501	-98.459
Adj. McFadden's R^2	0.417	0.390	0.444
Bayesian Info. Criteria	302.2	313.2	294.9

- The regression results are in $\frac{Odds\ Ratio\ ***}{(z\text{-}statistic)}$ format, where the Odds Ratio is the exponent of the standardized coefficient from the logistic regression.
† Significant at the 10% level; * Significant at the 5% level; **Significant at the 1% level; *** Significant at < 1% level.

To make the interpretation of the interaction terms easy, Model 9 was fitted again with unstandardized coefficients, and the regression coefficients for macroscale index at different values of microscale index and limited mode choice status were calculated (see Table 10). When there are limitations on mode choices and in a poor microscale environment, macroscale index actually had a negative coefficient (i.e., odds ratio less than one). However, even with limitations on mode choice options, a favorable microscale index led to a slightly positive coefficient for macroscale index (i.e., odds ratio of 1.056). Regardless of microscale index, trips with no limitations on mode choice had larger coefficients for macroscale index compared to those with the limitation. The coefficient of macroscale index was the largest when there are no limitations in mode choice and in a favorable microscale environment. When the mode choice limitation status is held constant, changing the microscale index value from 20th percentile to 80th percentile increases the macroscale index coefficients (i.e., logit) by about 0.37. When microscale index is held constant, switching from limited to not limited status leads to a roughly 0.72 increase in macroscale index coefficient.

Table 10. Coefficients for macroscale index at different values of household vehicle count and microscale index

Limited mode choice status	Microscale index	Effect of macroscale index	
		Logit	Odds ratio
Limited	8.0 (20th percentile)	-0.316	0.729
Limited	10.1 (80th percentile)	0.055	1.056
Not limited	8.0 (20th percentile)	0.403	1.496
Not limited	10.1 (80th percentile)	0.773	2.167

5.5 Discussion and conclusion

This chapter examines the association between walkability factors in all three scales of measurement (i.e., macro, meso, and microscales) and walking mode choice. It also focuses on two interaction terms that represent each of the two hypothetical pathways for the weaker relationship between disadvantaged populations and macroscale walkability factors. The interaction term corresponding to the first pathway – limited mode choice status as measured by vehicle ownership per adult and financial burden – showed that the effect of macroscale index on those with limitations on mode choice is weaker than it is on their counterpart. Similarly, the interaction term corresponding to the second pathway – concerns of safety, comfort, and pleasurability as measured by microscale factors – showed that the effect of macroscale index is weaker when coupled with poor microscale index. The difference in the model fit measures of Model 7 and 8 indicate that the first pathway may be the more prominent pathway among the two. Yet, the result of Model 9 suggests that both pathways are in effect simultaneously. Increasing microscale index from 20^{th} percentile to 80^{th} percentile and switching limited mode choice status are each linked with increases in macroscale coefficients (i.e., logit) by roughly 0.37 and 0.77, respectively.

This finding on the importance of microscale environment is particularly important for planners. Past studies that did not unpack the two pathways for the weaker relationship by examining their effect through separate variables are limited in providing guidance for intervention strategy from planning perspectives. This is because the variables used to characterize disadvantage were often conditions that planners cannot change, such as income and car ownership, race/ethnicity, and education. However, improving the microscale environment can be done using tools that planners can utilize. Moreover, microscale factors can be easier and faster to improve compared to macroscale factors. For example, increasing intersection density or land use mix of an established neighborhood by a meaningful amount can require significant redevelopments and can take multiple years. Such improvements on macroscale factors can be harder to be achieved in disadvantaged neighborhoods due to the fact that, with the exception of a few gentrifying neighborhoods, historically disinvested neighborhoods are seldom attractive investments for developers. In contrast, interventions on microscale factors, such as crosswalks, streetlights, and places to seat, can be easier and quicker to implement, and their benefits can be enjoyed instantly upon the installation.

Disadvantaged populations have a higher prevalence of unhealthy conditions associated with physical inactivity (Hardman & Stensel, 2009). This can be a significant public health challenge because disadvantaged populations' walking behavior tends to have weaker relationships with macroscale walkability factors. There are 96 trips made by individuals with limitations on mode choices, and about half of these trips were made in areas with poor microscale environment (i.e., lower than the median of microscale index), the combination that can suppress the effect of macroscale walkability factors. The findings of this chapter illustrate that encouraging disadvantaged populations with limitations on mode choice options to do more walking and other physical activities through macroscale-based interventions can have limited effectiveness unless it is coupled with improvements in microscale factors.

One important remaining challenge is the fact that microscale index also requires some level of macroscale index to have a positive association with walking mode choice. The interaction term consisting of macroscale and microscale indices in Model 9 indicates that the coefficient for microscale index can be zero when macroscale index is around -2.0, which is about 33th percentile. This suggests that some of the most promising opportunities for immediate interventions can be in areas that are moderate to high in macroscale index but poor in microscale environments. Figure 11 shows four quadrants defined by the median values of macro and microscale index, and it is quadrant C in Figure 11 that meets this condition. This quadrant is likely to benefit from the fact that microscale factors are easier and faster to improve and realize the full potential of the existing walking-supportive macroscale factors. Areas that are very low in macroscale index (i.e., quadrants A and D in Figure 11) may need interventions on both macroscale and microscale aspects of the built environment. Also note that the analysis of this chapter did not differentiate walking for transport and walking for recreation due to the limited sample size. Past studies suggested that microscale factors may be positively associated with recreational walking regardless of macroscale factors (Cain et al., 2014).

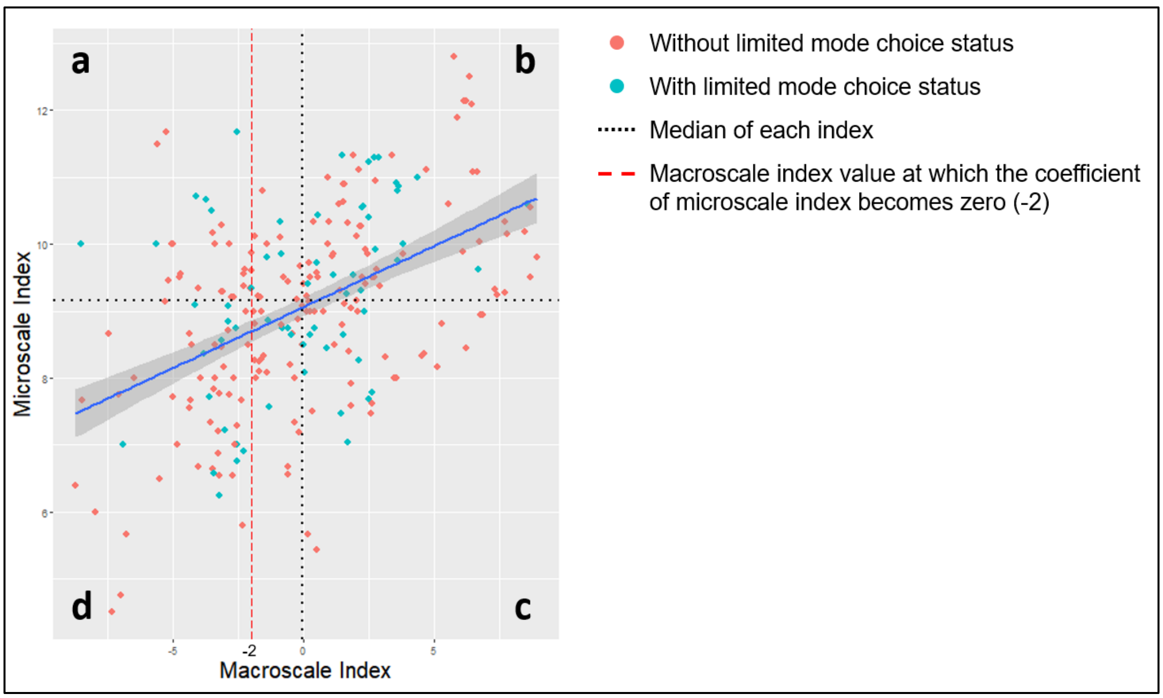


Figure 11. Scatterplot between macroscale and microscale index

The descriptive statistic in Table 6 and the t-test result show that the microscale environment is not statistically different between the activity space of those with and without limited mode choice status. This may seem to be in contrast to many past studies that reported inequitable distribution of microscale factors, such as Bereitschaft (2017). Below are a few potential explanations for this disparity. First, most of the past studies reporting the inequitable distribution of microscale factors examined the residential locations, whereas this chapter examined activity spaces that can contain other non-residential locations. Although activity spaces of different socioeconomic classes do not tend to overlap (Dong et al., 2020), the inclusion of non-residential locations may have diluted the differences between the two groups. Second, it may have been caused by the construction of the MAPS-mini, which include not only safety- and pleasurability-related qualities (e.g., graffities and building maintenance quality), which are likely to be better in wealthier subdivisions, but also pedestrian infrastructures (e.g., sidewalks and streetlights) which are more abundant in central cities. Considering that those with limited access to cars tend to concentrate in central cities due to their reliance on transit (Glaeser et al., 2008), people with limited mode choice are likely to have low scores on some MAPS-mini items while high scores for others. These inversely distributed items may have even out scores of microscale index between those with and without limited mode choice status to some degree. Third, the socio-demographic characteristics such as race, income, and educational attainment, which are used to characterize disadvantaged status in past studies, and limited mode choice status are related but not interchangeable concepts. For example, being a racial minority is not interchangeable with having limited mode choice options.

To demonstrate the plausibility of the three reasons above, t-test was conducted again with the following modifications: (1) the dataset is narrowed down to include only residential locations of each household, (2) only aesthetics-related items of the MAPS-mini are used to construct a modified microscale index (i.e., no graffiti, no ill-maintained buildings, presence of public park, presence of shade from overhead trees, and presence of seating), and (3) poverty status as determined by 2017 Federal Poverty Guidelines (U.S. Department of Health and Human Services, 2017) was used instead of limited mode choice status. This t-test under the modified setting showed that those that are at or under the poverty line are living in more unfavorable microscale environment (M: 1.710, SD: 0.508) than those above the poverty line (M:2.367, SD: 0.472), $t(9.041) = 3.426$, $p = 0.008$.

This chapter shares the research framework with Chapter 3, and many limitations applied to Chapter 3 also apply to this chapter: First, this study only considered origins of each trip due to the limited computational resources available to the author. Second, the design of the travel survey data introduced various limitations. The limited sample size prevented this chapter from subsetting the data by trip purpose or by home-based versus non-home-based trips. The sample size also may have resulted in the insignificance of some variables. Third, because the study site is limited to Atlanta, the degree to which the result of this chapter applies to other cities is unknown. Fourth, this chapter shares the potential sources of biases discussed in Chapter 3, including selective mobility and uncertain geographic context problem (Kwan, 2018).

In addition to these limitations, some limitations are unique to this chapter. First, there is a time difference between the year at which images used for microscale index were taken and the year the NHTS 2017 was conducted. The NHTS 2017 data was collected between 2016 and 2017, and most of the images for microscale index were taken between 2016 and 2020. If there were meaningful changes in the microscale environment between 2017 and 2020, which is not implausible considering the modifiability of microscale factors within a short time period, the coefficient estimates for microscale index would be biased. Second, the degree to which the quality of microscale factors of the built environment measured in this chapter relates to perceived quality is not known. Although both perceived and objectively measured environment variables associate with physical activity, they are likely to be distinct measures that "may capture different sources of variability in behavior" (Orstad et al., 2016, p. 917).

Additionally, the result of this chapter demonstrates that the automatically measured microscale factors are statistically significantly associated with walking mode choice, serving as additional validation for the automated audit method developed in the previous chapter. This chapter is an initial example of use cases of the automated audit method, which can be extended to future research in urban planning, transportation, public health, and other related fields.

CHAPTER 6. CONCLUSIONS

6.1 Summary

Past endeavors for measuring walkable built environment and empirical research using these measurements have disproportionately focused on macroscale factors of walkability. This is due to the difficulties in measuring meso- or microscale walkability factor, as their measurements often require manual labor and therefore tend to be labor- and resource-intensive. Although macroscale factors of walkability have proven to be highly effective in describing the walkable built environment, they are insufficient in presenting the comprehensive set of needs that pedestrians seek to fulfill when they make decisions to walk or not to walk. Pedestrians seek to fulfill various needs from the built environment as they walk, including the need for accessibility, safety, comfort, and pleasurability (M. A. Alfonzo, 2005). Accessibility is linked closely with macroscale factors (e.g., having places to go to and being physically connected to those places) while the other three higher-order needs can be closely proxied by walkability factors in smaller scales such as mesoscale and microscale (e.g., the quality of the experience going to places).

In short, the lack of a scalable method for measuring meso- and microscale factors of walkability has been a major challenge that resulted in (1) widely used walkability indices that fall short of reflecting various needs of pedestrians and (2) empirical research that puts disproportionate focus on macroscale factors of walkability, which are particularly common in the transportation field. Additionally, the non-existence of scalable auditing methods meant that urban planners and urban designers often lack street-level data on the microscale environment, which is required for identifying places in need of interventions.

This dissertation uses two recent technological advances to overcome these difficulties in measuring the built environment at meso and microscales and expands the literature by examining their association with walking mode choice. In Chapter 3, mesoscale factors of walkability were measured using street view images and a pre-trained computer vision model named Pyramid Scene Parsing Network. Three streetscape factors were crafted: building-to-street ratio, greenness, and sidewalk-to-street proportion. The chapter found that some streetscape factors are highly correlated with macroscale factors. A series of binary logistic regressions showed that streetscape factors (i.e., building-to-street ratio and greenness) are statistically significantly associated with walking mode choice. It was also found that streetscape factors contribute more to the fit of the models than macroscale factors.

Chapter 4 was devoted to the development and validation of a method for automatically auditing microscale factors. A computer vision architecture called Mask R-CNN was trained using a training dataset created by the author through transfer learning technique. The 15-item version of a validated walkability audit tool called the Microscale Audit of Pedestrian Streetscapes (MAPS-mini) was selected for automation. For validation of the performance of the automated audit method, interrater reliability was evaluated between the results from a human auditor and an automated audit method. The automated audit method produced kappa scores that are comparable to or better than the score from past studies that examined interrater reliability of human auditors. The total point, which is the sum of the scores of each microscale factor in the MAPS-mini, also showed a strong correlation between the human auditor and the automated audit.

Chapter 5 used the automated audit method from Chapter 4 to include microscale factors into the modeling framework used in Chapter 3. This chapter specifically focused on empirically unpacking the reasons for the weaker relationship between disadvantaged populations' walking behavior and macroscale walkability factors. The two hypothesized pathways for the weaker relationship – limited mode choice status and concerns for safety, comfort, and pleasurabilty – were proxied by the measure of vehicle availability per adult in households combined with financial burden and microscale index, respectively. Interaction terms consisting of macroscale index and the two proxy measures were added to the modeling strategy from Chapter 3. The results showed that the two hypothesized pathways are in effect simultaneously and that microscale index has a sizable moderating effect on macroscale index. The results can also be considered as an additional validation of the automated method developed in Chapter 4.

These results collectively demonstrate a promising outlook for using big image data such as GSV images and computer vision models in urban planning, public health, geography, and other related research.

6.2 Dissertation contributions

This dissertation offers theoretical and methodological contributes to the literature on the built environment and walking behavior relationship. Related fields of study that can benefit from the contributions include urban planning, transportation, public health, geography, and public policy.

6.2.1 Theoretical contributions

Because limitations in the current state of computer vision technology imposed restrictions on what aspect of the built environment can be measured, many past studies crafted walkability measures that were results of a compromise between what they hope to measure and what pre-trained computer vision models were able to measure. This compromise has led many studies to use indirect and remote proxies of walkability (e.g., sky view factor). Because of the indirectness, the theoretical link between the proxies and walking behavior was often unclear, and there has been a lack of an overarching theoretical framework through which these indirect proxies can be connected with the decision-making process of pedestrians.

This dissertation borrows the analytical approaches pertaining to the three different scales of measurements and merges them with the hypothesis on hierarchical walking needs to synthesize a comprehensive theoretical framework that can explain how various image-based measures of walkability can be connected with walking behavior (Figure 1). In this theoretical framework, macroscale factors of walkability tend to correspond with accessibility needs of pedestrians, which speaks to whether there are a variety of destinations to walk to in the vicinity of a given place and whether the destinations are physically and functionally connected with the trip origin. Mesoscale factors of walkability are visually perceivable size and scale of streetscapes which are linked with safety and comfort needs of pedestrians. Finally, microscale factors of walkability are the most fine-grained details and quality of streetscapes that relate to safety, comfort, and pleasurability needs.

This dissertation also adds to the literature on how the three scales of walkability factors affect one another with respect to their association with walking mode choice. Previous studies frequently theorized and modeled different scales of walkability factors as having additive effects on pedestrians' decision-making process. This dissertation illustrates a sizable moderating effect of microscale factors on macroscale factors and suggests that future studies can benefit from considering the interaction between walkability factors in different scales.

6.2.2 Methodological contribution

There are a few major methodological contributions in this dissertation. This dissertation is one of the newly emerging studies that combine street view images and computer vision technology in an attempt to automate walkability measurement.

Since around 2010, there have been attempts to use street view images for auditing walkable streetscapes. However, these early attempts were still conducting audits through human auditors (i.e., virtual audits), and the advantage in scalability was not substantial enough to apply virtual audit to larger geographic areas such as a city or a county. Recent advances in computer vision technology addressed the scalability issue by offering a way for automatically analyzing and extracting quantitative information from image data. However, as mentioned earlier, most pre-trained off-the-shelf computer vision models were limited to detecting a pre-defined set of objects, and these pre-defined objects seldom included the full suite of objects needed for comprehensive measures of walkable streetscapes.

In Chapter 3, this dissertation expands the literature by developing three mesoscale measures of walkable streetscapes. Methodological innovations specific to this chapter include (1) a systematic method for collecting street view images that combines the benefit of intersection-based and equidistance-based collection strategy and (2) three measures of walkable streetscapes that are consistent with the definition of mesoscale streetscapes.

The automated audit method developed in Chapter 4 is, to the best of the author's knowledge, one of the first attempts to automate the full suite of items of a scientifically validated walkability audit tool. Innovations specific to Chapter 4 include (1) a systematic method for collecting street view images that aims to download images that can cover as comprehensive aspects of any given streets as possible and (2) estimating and adjusting the gaps and/or overlaps in two consecutive street view images and automatically appending or cropping images.

6.3 Policy implications and suggestions for future research

6.3.1 Policy and planning implications

Two main contributions of this dissertation relevant to policy and planning implications are (1) the automated audit method and (2) the finding on the moderating effect of microscale factors on macroscale factors. With the automated audit method, urban planners and designers can start building databases on meso- and microscale environment at street-level over, for example, a city, a county, or even some larger geographies. This database can have the potential to allow urban planners and designers to provide targeted and timely interventions to streets in need of improvements.

Additionally, it would be ideal if these databases are built using street view images that are collected by local governments or similar public entities. The performance of the automated audit method is likely to be greatly improved if street view images were of higher quality. The lack of control over the exact location and time the images are taken, as well as geographic imbalance in the update interval, is another limitation that restricted the performance of the automated audit method. Image repositories created and maintained by public entities, such as local government or metropolitan planning organization, can have better control over these issues. These repositories can be invaluable sources of longitudinal, high-quality data, especially considering that the future advances in computer vision are expected to be able to extract even more accurate and diverse types of information from the images.

Urban planners and designers can make interventions in areas that have poor microscale factors mixed with some level of macroscale factors. Areas with very low levels of macroscale factors are not likely to benefit from interventions on microscale factors. When making these interventions, it is important to prioritize the activity space of disadvantaged populations with limited mode choice options, as those with limited mode choice options can be less responsive to, or have negative relationships with, the macroscale factors. This weak or negative relationship can be improved with a better microscale environment.

6.3.2 Suggestions for future research

In general, this dissertation demonstrates the usefulness of street view images as a data source for measuring walkable built environments. The results show that the street view image-based measurements can not only function as a proxy of macroscale factors but also provide additional information at eye level, which macroscale factors cannot capture.

More specifically, the limitations of the three analytical chapters call for future research in the following directions. First, future research is needed on how policies governing the built environment translate into streetscapes, particularly streetscapes in mesoscale (i.e., building-to-street ratio, greenness, sidewalk-to-street proportion). This translation is needed because such policies are often more relevant to the overhead-view measures, but the streetscape factors as presented in this chapter represent the streetscapes that arise as a result of the interplay of building height, the width of sidewalks and streets, and greenery in perspective view at eye level. This future research will benefit urban planners and designers by providing them with practical guidelines on, for example, how zoning and building codes, tree ordinances, and transportation plans can be leveraged to provide more walkable streetscapes. Provided that there exists a causal linkage between the built environment and walking behavior, this framework may provide practitioners a basis on which to articulate the expected outcome of their plans and designs.

Second, the computer vision models trained in Chapter 4 have much potential for improvements. Perhaps the most important factor in making the improvement is increasing the size of training data. Chapter 4 used about 4,000 images in total for the training of the three computer vision models, which is considerably smaller compared to other commonly used training image databases such as ADE20K, which contains roughly 25,000 training images and 2,000 validation images. Because creating a large enough training dataset with high quality is a labor-intensive task, one suggestion for the future endeavor is to create a coalition of researchers who collectively contribute to the creation of one training dataset that is designed specifically for the purpose of automated walkability audit.

Overall, this dissertation illustrated the effectiveness of measuring walkable built environment using large scale image data and using them in the built environment-walkability research. The future research can further calibrate the performance of the automated audit method as well as the applicability of the research findings to practice. Considering the near-global coverage of street view images, image-based walkability measurements can open new possibilities not only in data-rich regions such as the U.S. but also in regions where data constraints are a major barrier.

APPENDICES

A.1 Appendix – Chapter 4:

Table 11. Hyperparameters used in the transfer learning process

Hyperparameter	Segment model	Segment model (vertical)	Crossing model
BACKBONE	resnet101	resnet101	resnet101
BACKBONE_STRIDE	[4,8,16,32,64]	[4,8,16,32,64]	[4,8,16,32,64]
BATCH_SIZE	4	4	2
BBOX_STD_DEV	[0.1, 0.1, 0.2, 0.2]	[0.1, 0.1, 0.2, 0.2]	[0.1, 0.1, 0.2, 0.2]
COMPUTE_BACKBONE_SHAPE	None	None	None
DETECTION_MAX_INSTANCES	100	100	100
DETECTION_MIN_CONFIDENCE	0.7	0.7	0.7
DETECTION_NMS_THRESHOLD	0.3	0.3	0.3
FPN_CLASSIF_FC_LAYERS_SIZE	1024	1024	1024
GPU_COUNT	1	1	1
GRADIENT_CLIP_NORM	5	5	5
IMAGES_PER_GPU	4	4	2
IMAGE_CHANNEL_COUNT	3	3	3
IMAGE_MAX_DIM	1024	1024	1024
IMAGE_META_SIZE	33	15	22
IMAGE_MIN_DIM	640	640	640
IMAGE_MIN_SCALE	0	0	0
IMAGE_RESIZE_MODE	square	square	square
IMAGE_SHAPE	[1024 1024 3]	[1024 1024 3]	[1024 1024 3]
LEARNING_MOMENTUM	0.9	0.9	0.9
LEARNING_RATE	0.001	0.001	0.001
LOSS_WEIGHTS	rpn_class_loss:1, rpn_bbox_loss:1, mrcnn_class_loss:1, mrcnn_bbox_loss:1, mrcnn_mask_loss:1	rpn_class_loss:1, rpn_bbox_loss:1, mrcnn_class_loss:1, mrcnn_bbox_loss:1, mrcnn_mask_loss:1	rpn_class_loss:1, rpn_bbox_loss:1, mrcnn_class_loss:1, mrcnn_bbox_loss:1, mrcnn_mask_loss:1
MASK_POOL_SIZE	14	14	14
MASK_SHAPE	[28 28]	[28 28]	[28 28]
MAX_GT_INSTANCES	100	100	100
MINI_MASK_SHAPE	(56, 56)	(56, 56)	(56, 56)
NUM_CLASSES	21	3	10
POOL_SIZE	7	7	7
POST_NMS_ROIS_INFERENCE	1000	1000	1000
POST_NMS_ROIS_TRAINING	2000	2000	2000
PRE_NMS_LIMIT	6000	6000	6000
ROI_POSITIVE_RATIO	0.33	0.33	0.33
RPN_ANCHOR_RATIOS	[0.1, 1, 3]	[0.5, 1, 2]	[0.5, 1, 2]
RPN_ANCHOR_SCALES	(8, 32, 128, 256, 640)	(32, 64, 128, 256, 512)	(32, 64, 128, 256, 512)
RPN_ANCHOR_STRIDE	1	1	1
RPN_BBOX_STD_DEV	[0.1 0.1 0.2 0.2]	[0.1 0.1 0.2 0.2]	[0.1 0.1 0.2 0.2]
RPN_NMS_THRESHOLD	0.7	0.7	0.7
RPN_TRAIN_ANCHORS_PER_IMAGE	256	256	256
STEPS_PER_EPOCH	1000	1000	1000
TOP_DOWN_PYRAMID_SIZE	256	256	256
TRAIN_BN	FALSE	FALSE	FALSE
TRAIN_ROIS_PER_IMAGE	200	200	200
USE_MINI_MASK	TRUE	TRUE	TRUE
USE_RPN_ROIS	TRUE	TRUE	TRUE
VALIDATION_STEPS	50	50	50
WEIGHT_DECAY	0.00015	0.0001	0.0001

A.2 Appendix – Chapter 5:

Table 12. Results of the logistic regression models with microscale index in the form of average marginal effects (dep.var = walking / non-walking in binary)

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
Age	-0.055* (-0.10 to -0.01)	-0.078*** (-0.12 to -0.04)	-0.078*** (-0.12 to -0.04)	-0.056** (-0.1 to -0.01)	-0.067** (-0.11 to -0.02)	-0.050* (-0.09 to -0.01)
Gender – Male (Base: Female)	0.011 (-0.07 to 0.09)	0.039 (-0.04 to 0.12)	0.034 (-0.05 to 0.12)	0.009 (-0.07 to 0.09)	0.039 (-0.04 to 0.12)	0.015 (-0.06 to 0.09)
Race – Black (Base: White)	-0.060 (-0.17 to 0.05)	-0.005 (-0.11 to 0.1)	-0.032 (-0.14 to 0.08)	-0.057 (-0.17 to 0.05)	-0.013 (-0.12 to 0.09)	-0.039 (-0.14 to 0.07)
Race – Other races (Base: White)	0.013 (-0.14 to 0.17)	0.163* (0.03 to 0.29)	0.135* (0.00 to 0.27)	0.011 (-0.14 to 0.17)	0.161* (0.04 to 0.29)	0.040 (-0.11 to 0.19)
Household income	0.030 (-0.02 to 0.08)	0.057* (0.01 to 0.1)	0.064** (0.02 to 0.11)	0.030 (-0.02 to 0.08)	0.051* (0.01 to 0.1)	0.022 (-0.02 to 0.07)
Household vehicle count	-0.079** (-0.13 to -0.03)	-0.075** (-0.13 to -0.02)	-0.088** (-0.14 to -0.03)	-0.077** (-0.13 to -0.02)	-0.080** (-0.13 to -0.03)	-0.068** (-0.12 to -0.02)
Limited mode choice	-0.156** (-0.25 to -0.06)	-0.114* (-0.22 to -0.01)	-0.150** (-0.26 to -0.04)	-0.156** (-0.25 to -0.06)	-0.112* (-0.22 to -0.01)	-0.127* (-0.23 to -0.03)
Education – HS or higher (Base: Less than HS)	0.079 (-0.13 to 0.29)	0.256* (0.02 to 0.49)	0.175 (-0.07 to 0.42)	0.075 (-0.13 to 0.28)	0.246* (0.01 to 0.48)	0.153 (-0.06 to 0.37)
Driver status – Non-Driver (Base: Driver)	0.230*** (0.12 to 0.34)	0.200** (0.07 to 0.33)	0.189** (0.05 to 0.32)	0.228** (0.11 to 0.34)	0.219*** (0.09 to 0.35)	0.252*** (0.14 to 0.36)
Number walking activities in the past 7 days	0.111*** (0.07 to 0.15)	0.111*** (0.07 to 0.15)	0.122*** (0.08 to 0.17)	0.111*** (0.07 to 0.15)	0.106*** (0.06 to 0.15)	0.100*** (0.06 to 0.14)
Trip distance	-0.151*** (-0.18 to -0.12)	-0.190*** (-0.22 to -0.16)	-0.184*** (-0.22 to -0.15)	-0.151*** (-0.18 to -0.12)	-0.180*** (-0.21 to -0.15)	-0.152*** (-0.18 to -0.12)
Macroscale index	0.126*** (0.08 to 0.17)			0.131*** (0.08 to 0.18)		0.118*** (0.07 to 0.17)
Mesoscale index		0.068** (0.03 to 0.11)			0.076*** (0.04 to 0.12)	0.058** (0.02 to 0.1)
Microscale index			0.036† (-0.01 to 0.08)	-0.012 (-0.06 to 0.03)	0.049* (0.01 to 0.09)	0.003 (-0.04 to 0.05)
No. of observation	318	318	318	318	318	318
LL	-116.6	-125.6	-129.1	-116.5	-122.9	-112.9
Adj. McFadden's R^2	0.375	0.332	0.315	0.371	0.340	0.384
Bayesian Info. Criteria	308.2	326.1	333.2	313.7	326.5	312.2

- The regression results are in average marginal effects (confidence interval) format.

† Significant at the 10% level; * Significant at the 5% level; **Significant at the 1% level; *** Significant at < 1% level.

A.3 Appendix – Chapter 5:

Table 13. Results of the logistic regression models with microscale index in the form of average marginal effects (dep.var = walking / non-walking in binary)

	Model 7	Model 8	Model 9
Age	-0.038† (-0.08 to 0)	-0.055** (-0.1 to -0.01)	-0.044* (-0.08 to 0)
Gender – Male (Base: Female)	0.007 (-0.07 to 0.08)	0.026 (-0.05 to 0.1)	0.022 (-0.05 to 0.09)
Race – Black (Base: White)	-0.055 (-0.16 to 0.05)	-0.074 (-0.18 to 0.04)	-0.122* (-0.22 to -0.02)
Race – Other races (Base: White)	0.033 (-0.11 to 0.18)	0.019 (-0.13 to 0.17)	-0.002 (-0.14 to 0.14)
Household income	0.020 (-0.02 to 0.06)	0.021 (-0.03 to 0.07)	0.015 (-0.03 to 0.06)
Household vehicle count	-0.069** (-0.12 to -0.02)	-0.073** (-0.13 to -0.02)	-0.073** (-0.12 to -0.02)
Limited mode choice	-0.118* (-0.23 to -0.01)	-0.121* (-0.22 to -0.02)	-0.089 (-0.20 to 0.02)
Education – HS or higher (Base: Less than HS)	0.148 (-0.06 to 0.36)	0.137 (-0.06 to 0.34)	0.165† (-0.01 to 0.34)
Driver status – Non-Driver (Base: Driver)	0.219*** (0.10 to 0.34)	0.239*** (0.13 to 0.35)	0.206*** (0.09 to 0.32)
Number walking activities in the past 7 days	0.088*** (0.05 to 0.12)	0.101*** (0.06 to 0.14)	0.083*** (0.05 to 0.12)
Trip distance	-0.146*** (-0.18 to -0.11)	-0.151*** (-0.18 to -0.12)	-0.146*** (-0.18 to -0.11)
Macroscale index	0.111*** (0.06 to 0.16)	0.114*** (0.06 to 0.17)	0.102*** (0.05 to 0.15)
Mesoscale index	0.066** (0.03 to 0.11)	0.052* (0.01 to 0.09)	0.059** (0.02 to 0.1)
Microscale index	0.002 (-0.04 to 0.05)	0.011 (-0.03 to 0.06)	0.019 (-0.03 to 0.06)
No. of observation	318	318	318
LL	-105.006	-110.501	-98.459
Adj. McFadden's R^2	0.417	0.390	0.444
Bayesian Info. Criteria	302.2	313.2	294.9

- The regression results are in average marginal effects (confidence interval) format.
† Significant at the 10% level; * Significant at the 5% level; **Significant at the 1% level; *** Significant at < 1% level.

REFERENCES

Adkins, A., Dill, J., Luhr, G., & Neal, M. (2012). Unpacking Walkability: Testing the Influence of Urban Design Features on Perceptions of Walking Environment Attractiveness. *Journal of Urban Design*, *17*(4), 499–510. https://doi.org/10.1080/13574809.2012.706365

Adkins, A., Makarewicz, C., Scanze, M., Ingram, M., & Luhr, G. (2017). Contextualizing Walkability: Do Relationships Between Built Environments and Walking Vary by Socioeconomic Context? *Journal of the American Planning Association*, *83*(3), 296–314. https://doi.org/10.1080/01944363.2017.1322527

Alfonzo, M. A. (2005). To Walk or Not to Walk? The Hierarchy of Walking Needs. *Environment and Behavior*, *37*(6), 808–836. https://doi.org/10.1177/0013916504274016

Alfonzo, M., Boarnet, M. G., Day, K., Mcmillan, T., & Anderson, C. L. (2008). The Relationship of Neighbourhood Built Environment Features and Adult Parents' Walking. *Journal of Urban Design*, *13*(1), 29–51. https://doi.org/10.1080/13574800701803456

Bader, M. D. M., Mooney, S. J., Bennett, B., & Rundle, A. G. (2017). The Promise, Practicalities, and Perils of Virtually Auditing Neighborhoods Using Google Street View. *The ANNALS of the American Academy of Political and Social Science*, *669*(1), 18–40. https://doi.org/10.1177/0002716216681488

Badland, H. M., Opit, S., Witten, K., Kearns, R. A., & Mavoa, S. (2010). Can Virtual Streetscape Audits Reliably Replace Physical Streetscape Audits? *Journal of Urban Health*, *87*(6), 1007–1016. https://doi.org/10.1007/s11524-010-9505-x

Bennett, G. G., McNeill, L. H., Wolin, K. Y., Duncan, D. T., Puleo, E., & Emmons, K. M. (2007). Safe To Walk? Neighborhood Safety and Physical Activity Among Public Housing Residents. *PLOS Medicine*, *4*(10), e306. https://doi.org/10.1371/journal.pmed.0040306

Bereitschaft, B. (2017). Equity in microscale urban design and walkability: A photographic survey of six Pittsburgh streetscapes. *Sustainability (Switzerland)*, *9*(7), 1233. https://doi.org/10.3390/su9071233

Borst, H. C., Miedema, H. M. E., de Vries, S. I., Graham, J. M. A., & van Dongen, J. E. F. (2008). Relationships between street characteristics and perceived attractiveness for walking reported by elderly people. *Journal of Environmental Psychology*, *28*(4), 353–361. https://doi.org/10.1016/j.jenvp.2008.02.010

Brookfield, K., & Tilley, S. (2016). Using Virtual Street Audits to Understand the Walkability of Older Adults' Route Choices by Gender and Age. *International Journal of Environmental Research and Public Health*, *13*(11). https://doi.org/10.3390/ijerph13111061

Cain, K. L., Millstein, R. A., Sallis, J. F., Conway, T. L., Gavand, K. A., Frank, L. D., Saelens, B. E., Geremia, C. M., Chapman, J., Adams, M. A., Glanz, K., & King, A. C. (2014). Contribution of streetscape audits to explanation of physical activity in four age groups based on the Microscale Audit of Pedestrian Streetscapes (MAPS). *Social Science & Medicine*, *116*, 82–92. https://doi.org/10.1016/j.socscimed.2014.06.042

Cervero, R., & Duncan, M. (2003). Walking, bicycling, and urban landscapes: Evidence from the San Francisco Bay Area. *American Journal of Public Health*, *93*(9), 1478–1483. https://doi.org/10.2105/ajph.93.9.1478

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Cervero, R., Sarmiento, O. L., Jacoby, E., Gomez, L. F., & Neiman, A. (2009). Influences of Built Environments on Walking and Cycling: Lessons from Bogotá. *International Journal of Sustainable Transportation*, *3*(4), 203–226. https://doi.org/10.1080/15568310802178314

Chichester, B. M., Gregan, J. A., Anderson, D. P., & Kerr, J. M. (1998). Associations between Road Traffic Accidents and Socio-Economic Deprivation on Scotland's West Coast. *Scottish Medical Journal*, *43*(5), 135–138. https://doi.org/10.1177/003693309804300503

Chiu, M., Shah, B. R., Maclagan, L. C., Rezai, M.-R., Austin, P. C., & Tu, J. V. (2015). Walk Score® and the prevalence of utilitarian walking and obesity among Ontario adults: A cross-sectional study. *Health Reports*, *26*(7), 3–10.

Clarke, P., Ailshire, J., Melendez, R., Bader, M., & Morenoff, J. (2010). Using Google Earth to conduct a neighborhood audit: Reliability of a virtual audit instrument. *Health & Place*, *16*(6), 1224–1229. https://doi.org/10.1016/J.HEALTHPLACE.2010.08.007

Clifton, K. J., Livi Smith, A. D., & Rodriguez, D. (2007). The development and testing of an audit for the pedestrian environment. *Landscape and Urban Planning*, *80*(1), 95–110. https://doi.org/10.1016/j.landurbplan.2006.06.008

Day, K. (2006). Active Living and Social Justice: Planning for Physical Activity in Low-income, Black, and Latino Communities. *Journal of the American Planning Association*, *72*(1), 88–99. https://doi.org/10.1080/01944360608976726

Dong, X., Morales, A. J., Jahani, E., Moro, E., Lepri, B., Bozkaya, B., Sarraute, C., Bar-Yam, Y., & Pentland, A. (2020). Segregated interactions in urban and online space. *EPJ Data Sci.*, *9*(1). https://doi.org/10.1140/epjds/s13688-020-00238-7

Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016). Deep Learning the City: Quantifying Urban Perception at a Global Scale. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer Vision – ECCV 2016* (pp. 196–212). Springer International Publishing.

Duncan, D. T., Aldstadt, J., Whalen, J., Melly, S. J., Gortmaker, S. L., Duncan, D. T., Aldstadt, J., Whalen, J., Melly, S. J., & Gortmaker, S. L. (2011). Validation of Walk Score® for Estimating Neighborhood Walkability: An Analysis of Four US Metropolitan Areas. *International Journal of Environmental Research and Public Health*, *8*(11), 4160–4179. https://doi.org/10.3390/ijerph8114160

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Ewing, R., & Clemente, Otto. (2013). *Measuring urban design: Metrics for livable places*. Island Press.

Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban Design*, *14*(1), 65–84. https://doi.org/10.1080/13574800802451155

Foltête, J.-C., & Piombini, A. (2007). Urban layout, landscape features and pedestrian usage. *Landscape and Urban Planning*, *81*(3), 225–234. https://doi.org/10.1016/J.LANDURBPLAN.2006.12.001

Forsyth, A., Michael Oakes, J., Lee, B., & Schmitz, K. H. (2009). The built environment, walking, and physical activity: Is the environment more important to some people than others? *Transportation Research Part D: Transport and Environment*, *14*(1), 42–49. https://doi.org/10.1016/j.trd.2008.10.003

Frank, L. D., Andresen, M. A., & Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine*, *27*(2), 87–96. https://doi.org/10.1016/j.amepre.2004.04.011

Frank, L. D., Kerr, J., Sallis, J. F., Miles, R., & Chapman, J. (2008). A hierarchy of sociodemographic and environmental correlates of walking and obesity. *Preventive Medicine*, *47*(2), 172–178. https://doi.org/10.1016/j.ypmed.2008.04.004

Frank, L. D., Schmid, T. L., Sallis, J. F., Chapman, J., & Saelens, B. E. (2005). Linking objectively measured physical activity with objectively measured urban form: Findings from SMARTRAQ. *American Journal of Preventive Medicine*, *28*, 117–125. https://doi.org/10.1016/j.amepre.2004.11.001

Fry, D., Mooney, S. J., Rodríguez, D. A., Caiaffa, W. T., & Lovasi, G. S. (2020). Assessing Google Street View Image Availability in Latin American Cities. *Journal of Urban Health*, *97*(4), 552–560. https://doi.org/10.1007/s11524-019-00408-7

Gallimore, J. M., Brown, B. B., & Werner, C. M. (2011). Walking routes to school in new urban and suburban neighborhoods: An environmental walkability analysis of blocks and routes. *Journal of Environmental Psychology*, *31*(2), 184–191. https://doi.org/10.1016/J.JENVP.2011.01.001

Giarrusso, A. J., & Smith, S. M. (2014). *Assessing Urban Tree Canopy in the City of Atlanta: A Baseline Canopy Study*. Georgia Institute of Technology.

Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, *63*(1), 1–24. https://doi.org/10.1016/j.jue.2006.12.004

Glaeser, E. L., Kominers, S. D., Luca, M., & Naik, N. (2018). Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life. *Economic Inquiry*, *56*(1), 114–137. https://doi.org/10.1111/ecin.12364

Griew, P., Hillsdon, M., Foster, C., Coombes, E., Jones, A., & Wilkinson, P. (2013). Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, *10*(1), 103. https://doi.org/10.1186/1479-5868-10-103

Gullón, P., Badland, H. M., Alfayate, S., Bilal, U., Escobar, F., Cebrecos, A., Diez, J., & Franco, M. (2015). Assessing Walking and Cycling Environments in the Streets of Madrid: Comparing On-Field and Virtual Audits. *Journal of Urban Health*, *92*(5), 923–939. https://doi.org/10.1007/s11524-015-9982-z

Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity. *American Journal of Preventive Medicine*, *23*(2), 64–73. https://doi.org/10.1016/s0749-3797(02)00475-0

Hankey, S., Zhang, W., Le, H. T. K., Hystad, P., & James, P. (2021). Predicting bicycling and walking traffic using street view imagery and destination data. *Transportation Research Part D: Transport and Environment*, *90*, 102651. https://doi.org/10.1016/j.trd.2020.102651

Hanson, C. S., Noland, R. B., & Brown, C. (2013). The severity of pedestrian crashes: An analysis using Google Street View imagery. *Journal of Transport Geography*, *33*, 42–53. https://doi.org/10.1016/j.jtrangeo.2013.09.002

Hardman, A. E., & Stensel, D. J. (2009). *Physical acitivty and health, The evidence explained*. Routledge.

Harrison, R. A., Gemmell, I., & Heller, R. F. (2007). The population effect of crime and neighbourhood on physical activity: An analysis of 15 461 adults. *Journal of Epidemiology and Community Health*, *61*(1), 34. https://doi.org/10.1136/jech.2006.048389

Harvey, C., & Aultman-Hall, L. (2015). Urban Streetscape Design and Crash Severity. *Transportation Research Record: Journal of the Transportation Research Board*, *2500*(1), 1–8. https://doi.org/10.3141/2500-01

Harvey, C., & Aultman-Hall, L. (2016). Measuring Urban Streetscapes for Livability: A Review of Approaches. *The Professional Geographer*, *68*(1), 149–158. https://doi.org/10.1080/00330124.2015.1065546

Harvey, C., Aultman-Hall, L., Hurley, S. E., & Troy, A. (2015). Effects of skeletal streetscape design on perceived safety. *Landscape and Urban Planning*, *142*, 18–28. https://doi.org/10.1016/j.landurbplan.2015.05.007

Harvey, C., Aultman-Hall, L., Troy, A., & Hurley, S. E. (2017). Streetscape skeleton measurement and classification. *Environment and Planning B: Urban Analytics and City Science*, *44*(4), 668–692. https://doi.org/10.1177/0265813515624688

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. *IEEE International Conference on Computer Vision*, 2961–2969.

http://openaccess.thecvf.com/content_ICCV_2017/papers/He_Mask_R-CNN_ICCV_2017_paper.pdf

He, L., Páez, A., & Liu, D. (2017). Built environment and violent crime: An environmental audit approach using Google Street View. *Computers, Environment and Urban Systems*, *66*, 83–95. https://doi.org/10.1016/J.COMPENVURBSYS.2017.08.001

Hipp, J. A., Adlakha, D., Eyler, A. A., Chang, B., & Pless, R. (2013). Emerging Technologies: Webcams and Crowd-Sourcing to Identify Active Transportation. *American Journal of Preventive Medicine*, *44*(1), 96–97. https://doi.org/10.1016/j.amepre.2012.09.051

Hooker, S. P., Wilson, D. K., Griffin, S. F., & Ainsworth, B. E. (2005). Perceptions of Environmental Supports for Physical Activity in African American and White Adults in a Rural County in South Carolina. *Prev Chronic Dis. 2005; 2(4).*, *2*(4). https://stacks.cdc.gov/view/cdc/19964

Huston, S. L., Evenson, K. R., Bors, P., & Gizlice, Z. (2003). Neighborhood Environment, Access to Places for Activity, and Leisure-Time Physical Activity in a Diverse North Carolina Population. *American Journal of Health Promotion*, *18*(1), 58–69. https://doi.org/10.4278/0890-1171-18.1.58

Hwang, J., & Sampson, R. J. (2014). Divergent Pathways of Gentrification: Racial Inequality and the Social Order of Renewal in Chicago Neighborhoods. *American Sociological Review*, *79*(4), 726–751. https://doi.org/10.1177/0003122414535774

Ivory, V. C., Blakely, T., Pearce, J., Witten, K., Bagheri, N., Badland, H., & Schofield, G. (2015). Could strength of exposure to the residential neighbourhood modify associations between walkability and physical activity? *Social Science & Medicine*, *147*, 232–241. https://doi.org/10.1016/j.socscimed.2015.10.053

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790. https://doi.org/10.1126/science.aaf7894

Jiang, B., Deal, B., Pan, H., Larsen, L., Hsieh, C.-H., Chang, C.-Y., & Sullivan, W. C. (2017). Remotely-sensed imagery vs. Eye-level photography: Evaluating associations among measurements of tree cover density. *Landscape and Urban Planning*, *157*, 270–281. https://doi.org/10.1016/J.LANDURBPLAN.2016.07.010

Kelly, C. M., Wilson, J. S., Baker, E. A., Miller, D. K., & Schootman, M. (2013). Using Google Street View to Audit the Built Environment: Inter-rater Reliability

Results. *Annals of Behavioral Medicine*, *45*(S1), 108–112. https://doi.org/10.1007/s12160-012-9419-9

Kelly, C., Wilson, J. S., Schootman, M., Clennin, M., Baker, E. A., & Miller, D. K. (2014). The Built Environment Predicts Observed Physical Activity. *Frontiers in Public Health*, *2*, 52. https://doi.org/10.3389/fpubh.2014.00052

Kerr, J., Frank, L., Sallis, J. F., & Chapman, J. (2007). Urban form correlates of pedestrian travel in youth: Differences by gender, race-ethnicity and household attributes. *Transportation Research Part D: Transport and Environment*, *12*(3), 177–182. https://doi.org/10.1016/j.trd.2007.01.006

Ki, D., & Lee, S. (2021). Analyzing the effects of Green View Index of neighborhood streets on walking time using Google Street View and deep learning. *Landscape and Urban Planning*, *205*, 103920. https://doi.org/10.1016/j.landurbplan.2020.103920

King, W., & Palmisano, P. (1992). Racial differences in childhood hospitalized pedestrian injuries. *Pediatric Emergency Care*, *8*(4), 221–224. PubMed. https://doi.org/10.1097/00006565-199208000-00013

Koo, B. W., Boyd, N., Botchwey, N., & Guhathakurta, S. (2019). Environmental Equity and Spatiotemporal Patterns of Urban Tree Canopy in Atlanta. *Journal of Planning Education and Research*, 0739456X19864149. https://doi.org/10.1177/0739456X19864149

Kwan, M. P. (2018). The Limits of the Neighborhood Effect: Contextual Uncertainties in Geographic, Environmental Health, and Social Science Research. *Annals of the American Association of Geographers*, *108*(6), 1482–1490. https://doi.org/10.1080/24694452.2018.1453777

Larkin, A., & Hystad, P. (2019). Evaluating street view exposure measures of visible green space for health research. *Journal of Exposure Science & Environmental Epidemiology*, *29*(4), 447–456. https://doi.org/10.1038/s41370-018-0017-1

Li, X., Ratti, C., & Seiferling, I. (2018). Quantifying the shade provision of street trees in urban landscape: A case study in Boston, USA, using Google Street View. *Landscape and Urban Planning*, *169*, 81–91. https://doi.org/10.1016/j.landurbplan.2017.08.011

Li, X., Santi, P., Courtney, T. K., Verma, S. K., & Ratti, C. (2018). Investigating the association between streetscapes and human walking activities using Google

Street View and human trajectory data. *Transactions in GIS*, *22*(4), 1029–1044. https://doi.org/10.1111/tgis.12472

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, *14*(3), 675–685. https://doi.org/10.1016/J.UFUG.2015.06.006

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740–755). Springer International Publishing.

Long, J., Shelhamer, E., & Darrell, T. (2015). *Fully Convolutional Networks for Semantic Segmentation*. 3431–3440.

Lovasi, G. S., Hutson, M. A., Guerra, M., & Neckerman, K. M. (2009). Built Environments and Obesity in Disadvantaged Populations. *Epidemiologic Reviews*, *31*(1), 7–20. https://doi.org/10.1093/epirev/mxp005

Lovasi, G. S., Neckerman, K. M., Quinn, J. W., Weiss, C. C., & Rundle, A. (2009). Effect of individual or neighborhood disadvantage on the association between neighborhood walkability and body mass index. *American Journal of Public Health*, *99*(2), 279–284. https://doi.org/10.2105/AJPH.2008.138230

Lu, Y., Yang, Y., Sun, G., & Gou, Z. (2019). Associations between overhead-view and eye-level urban greenness and cycling behaviors. *Cities*, *88*, 10–18. https://doi.org/10.1016/J.CITIES.2019.01.003

Manaugh, K., & El-Geneidy, A. (2011). Validating walkability indices: How do different households respond to the walkability of their neighborhood? *Transportation Research Part D: Transport and Environment*, *16*(4), 309–315. https://doi.org/10.1016/J.TRD.2011.01.009

Marco, M., Gracia, E., Martín-Fernández, M., & López-Quílez, A. (2017). Validation of a Google Street View-Based Neighborhood Disorder Observational Scale. *Journal of Urban Health : Bulletin of the New York Academy of Medicine*, *94*(2), 190–198. https://doi.org/10.1007/s11524-017-0134-5

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. PubMed.

Mertens, L., Van Cauwenberg, J., Ghekiere, A., Van Holle, V., De Bourdeaudhuij, I., Deforche, B., Nasar, J., Van de Weghe, N., & Van Dyck, D. (2015). Does the Effect of Micro-Environmental Factors on a Street's Appeal for Adults' Bicycle Transport Vary across Different Macro-Environments? An Experimental Study. *PLOS ONE*, *10*(8), e0136715. https://doi.org/10.1371/journal.pone.0136715

Mooney, S. J., DiMaggio, C. J., Lovasi, G. S., Neckerman, K. M., Bader, M. D. M., Teitler, J. O., Sheehan, D. M., Jack, D. W., & Rundle, A. G. (2016). Use of Google Street View to Assess Environmental Contributions to Pedestrian Injury. *American Journal of Public Health*, *106*(3), 462–469. https://doi.org/10.2105/AJPH.2015.302978

Neckerman, K. M., Lovasi, G. S., Davies, S., Purciel, M., Quinn, J., Feder, E., Raghunath, N., Wasserman, B., & Rundle, A. (2009). Disparities in Urban Neighborhood Conditions: Evidence from GIS Measures and Field Observation in New York City. *Journal of Public Health Policy*, *30*(S1), S264–S285. https://doi.org/10.1057/jphp.2008.47

Nguyen, Q. C., Khanna, S., Dwivedi, P., Huang, D., Huang, Y., Tasdizen, T., Brunisholz, K. D., Li, F., Gorman, W., Nguyen, T. T., & Jiang, C. (2019). Using Google Street View to examine associations between built environment characteristics and U.S. health outcomes. *Preventive Medicine Reports*, *14*, 100859. https://doi.org/10.1016/J.PMEDR.2019.100859

Odgers, C. L., Caspi, A., Bates, C. J., Sampson, R. J., & Moffitt, T. E. (2012). Systematic social observation of children's neighborhoods using Google Street View: A reliable and cost-effective method. *Journal of Child Psychology and Psychiatry*, *53*(10), 1009–1017. https://doi.org/10.1111/j.1469-7610.2012.02565.x

Orstad, S. L., McDonough, M. H., Stapleton, S., Altincekic, C., & Troped, P. J. (2016). A Systematic Review of Agreement Between Perceived and Objective Neighborhood Environment Measures and Associations With Physical Activity Outcomes. *Environment and Behavior, 49*(8), 904–932. https://doi.org/10.1177/0013916516670982

Park, S., Choi, K., & Lee, J. S. (2015). To Walk or Not to Walk: Testing the Effect of Path Walkability on Transit Users' Access Mode Choices to the Station. *International Journal of Sustainable Transportation*, *9*(8), 529–541. https://doi.org/10.1080/15568318.2013.825036

Rundle, A. G., Bader, M. D. M., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to Audit Neighborhood Environments.

American Journal of Preventive Medicine, *40*(1), 94–100. https://doi.org/10.1016/j.amepre.2010.09.034

Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine*, *25*(2), 80–91. https://doi.org/10.1207/S15324796ABM2502_03

Sallis, J. F., Cain, K. L., Conway, T. L., Gavand, K. A., Millstein, R. A., Geremia, C. M., Frank, L. D., Saelens, B. E., Glanz, K., & King, A. C. (2015). Is Your Neighborhood Designed to Support Physical Activity? A Brief Streetscape Audit Tool. *Preventing Chronic Disease*, *12*. https://doi.org/10.5888/pcd12.150098

Sallis, J. F., Floyd, M. F., Rodríguez, D. A., & Saelens, B. E. (2012). Role of Built Environments in Physical Activity, Obesity, and Cardiovascular Disease. *Circulation*, *125*(5), 729–737. https://doi.org/10.1161/CIRCULATIONAHA.110.969022

Sallis, J. F., Slymen, D. J., Conway, T. L., Frank, L. D., Saelens, B. E., Cain, K., & Chapman, J. E. (2011). Income disparities in perceived neighborhood built and social environment attributes. *Health & Place*, *17*(6), 1274–1283. https://doi.org/10.1016/J.HEALTHPLACE.2011.02.006

Seiferling, I., Naik, N., Ratti, C., & Proulx, R. (2017). Green streets − Quantifying and mapping urban trees with street-level imagery and computer vision. *Landscape and Urban Planning*, *165*, 93–101. https://doi.org/10.1016/J.LANDURBPLAN.2017.05.010

Sim, J., & Wright, C. C. (2005). The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*, *85*(3), 257–268. https://doi.org/10.1093/ptj/85.3.257

Smith, K. R., Brown, B. B., Yamada, I., Kowaleski-Jones, L., Zick, C. D., & Fan, J. X. (2008). Walkability and Body Mass Index: Density, Design, and New Diversity Measures. *American Journal of Preventive Medicine*, *35*(3), 237–244. https://doi.org/10.1016/J.AMEPRE.2008.05.028

Tang, J., & Long, Y. (2018). Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing. *Landscape and Urban Planning*. https://doi.org/10.1016/J.LANDURBPLAN.2018.09.015

U.S. Department of Health and Human Services. (2017). *2017 Federal Poverty Guidelines*. https://dch.georgia.gov/federal-poverty-guidelines-0

U.S. Environmental Protection Agency. (2015). *National Walkability Index*. https://www.epa.gov/smartgrowth/smart-location-mapping

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, *2018*, 7068349. https://doi.org/10.1155/2018/7068349

Walk Score. (n.d.). *Walk Score Methodology*. https://www.walkscore.com/methodologyhtml

Wang, R., Helbich, M., Yao, Y., Zhang, J., Liu, P., Yuana, Y., & Liu, Y. (2019). *Urban greenery and mental wellbeing in adults: Cross-sectional mediation analyses on multiple pathways across different greenery measures*. http://arxiv.org/abs/1905.04488

Wang, R., Liu, Y., Lu, Y., Yuan, Y., Zhang, J., Liu, P., & Yao, Y. (2019). The linkage between the perception of neighbourhood and physical activity in Guangzhou, China: Using street view imagery with deep learning techniques. *International Journal of Health Geographics*, *18*(1), 18. https://doi.org/10.1186/s12942-019-0182-z

Wang, R., Lu, Y., Zhang, J., Liu, P., Yao, Y., & Liu, Y. (2019). The relationship between visual enclosure for neighbourhood street walkability and elders' mental health in China: Using street view images. *Journal of Transport & Health*, *13*, 90–102. https://doi.org/10.1016/J.JTH.2019.02.009

Weisel, D. L. (2002). *Graffiti* (Guide No.9). Arizona State University Center for Problem-Oriented Policing. https://popcenter.asu.edu/content/graffiti-0

Yin, L., Cheng, Q., Wang, Z., & Shao, Z. (2015). 'Big data' for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts. *Applied Geography*, *63*, 337–345. https://doi.org/10.1016/j.apgeog.2015.07.010

Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. *Applied Geography*, *76*, 147–153. https://doi.org/10.1016/J.APGEOG.2016.09.024

Zhang, L., Ye, Y., Zeng, W., Chiaradia, A., Zhang, L., Ye, Y., Zeng, W., & Chiaradia, A. (2019). A Systematic Measurement of Street Quality through Multi-Sourced Urban Data: A Human-Oriented Analysis. *International Journal of*

Environmental Research and Public Health, *16*(10), 1782. https://doi.org/10.3390/ijerph16101782

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). *Pyramid Scene Parsing Network*. 2881–2890.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. (2017). Scene Parsing through ADE20K Dataset. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 633–641. http://groups.csail.mit.edu/vision/datasets/ADE20K/

Zhu, X., & Lee, C. (2008). Walkability and Safety Around Elementary Schools: Economic and Ethnic Disparities. *American Journal of Preventive Medicine*, *34*(4), 282–290. https://doi.org/10.1016/J.AMEPRE.2008.01.024